ISSN 1330-3651(Print), ISSN 1848-6339 (Online)
Received: 7 May 2025; Accepted: 20 November 2025

https://doi.org/10.17559/TV-20250507002651
Original scientific paper

Estimation of Rock Brittleness from Point Load Strength Index Data Using Machine Learning Methods

Deniz AKBAY*, Gökhan EKINCIOGLU, Murat ISIK, Mehmet Ali YALCINKAYA

Abstract: Brittleness is a vital mechanical property that characterizes a rock's tendency to fracture under applied stress without significant deformation, which is particularly significant in mining, tunnelling, and other geotechnical engineering applications. The accurate prediction of rock brittleness is essential for optimizing excavation strategies, ensuring operational safety, and improving the cost-efficiency of resource extraction processes. However, conventional brittleness assessment techniques-such as those based on uniaxial compressive strength (UCS) and tensile strength-can be labour-intensive, time-consuming, and expensive. This study introduces a predictive framework based on machine learning algorithms using Point Load Strength Index (PLI) values as the sole input variable. A comprehensive dataset comprising sedimentary, igneous, and metamorphic rocks was compiled from both literature sources and laboratory experiments. Multiple regression models were applied and compared, including traditional linear methods and advanced ensemble learners. Among these, the Gradient Boosting Regressor delivered the highest predictive accuracy, achieving an (R^2) value of 0.96 for metamorphic rocks. The results demonstrate that even a single indirect measurement like PLI can serve as an effective predictor of rock brittleness when coupled with robust machine learning techniques. The findings highlight the potential of integrating AI-based models into rock mechanics workflows to streamline brittleness estimation and support sustainable mining practices.

Keywords: geotechnical engineering; machine learning; non-destructive testing; point load strength index; rock brittleness

1 INTRODUCTION

In rock mechanics and geotechnical applications, brittleness is a crucial mechanical parameter that defines a material's sensitivity to sudden failure without significant deformation. Several researchers have so far put forth alternative ways to measure rock brittleness based on various ideas, taking into account a variety of influencing elements such as strength parameters, in-situ stress, and mineral composition. Each index's dependability depends on whether the right strategy is applied for the intended use [1].

Brittleness is the ability of a material to continuously deform, fracture, or crack under force without undergoing permanent deformations prior to failure [2].

The brittleness of rock significantly influences its mechanical behavior and failure characteristics, particularly in applications involving rock excavation, such as drillability and cuttability [3-7], fracability during hydraulic fracturing [8-13], and rockburst proneness in deep hard rock tunnels [14]. Reliable evaluation of rock brittleness is essential for several critical engineering tasks, including the appropriate design and selection of excavation and mining equipment, the stimulation of unconventional shale gas formations, and the stability analysis of deep hard rock tunnels.

The brittleness values of rocks are typically determined in engineering applications using the tensile strength (TS) and uniaxial compressive strength (UCS) of the rock [6, 14-23]. The motivation of the study is that conventional brittleness assessment based on UCS and TS requires specimen preparation and specialized equipment, making testing costly and time-consuming. In practice, indirect tests such as the Point Load Strength Index (PLI) are favoured because they are simpler and faster. The research question of the study is whether the B_3 brittleness index of intact rocks can be accurately predicted from PLI values alone using machine-learning models, and which model provides the most reliable performance across sedimentary, igneous, and metamorphic rocks.

This study makes three practical contributions. First, we curate and harmonize a cross-study dataset of 520 rock samples reporting PLI and brittleness indices (B_1-B_4); assembling such breadth from a single experimental campaign would be prohibitively costly and time-consuming. Second, we provide an apples-to-apples benchmark of 15 representative regression families for the single-predictor PLI→B_3 problem under a uniform 5-fold CV protocol, clarifying which families are most reliable across lithologies. Third, by evaluating 15 algorithms spanning complementary families - linear and regularized (Linear, Ridge, Lasso, Elastic Net, Bayesian Ridge), robust linear (Huber), kernel/instance-based (SVR, KNN), tree-based and ensembles (Decision Tree, Random Forest, Extra Trees, Gradient Boosting, AdaBoost), non-parametric Bayesian (Gaussian Process), and neural (MLP) we can directly examine how different inductive biases perform on the PLI→B_3 task and compare their effectiveness across lithologies.

2 RELATED WORKS

Over the past five decades, researchers have developed a wide range of brittleness indices tailored to different applications within rock mechanics [1, 15, 24, 25]. The concept and measurement of rock brittleness have been widely varied and have not yet been standardised, despite the fact that brittleness indices are widely used in various rock engineering sectors. Additionally, the majority of the indices were developed to serve certain application areas [24].

In recent years, numerous studies have been conducted to define rock brittleness. In addition, several studies have reported strong correlations between rock brittleness and various elastic properties, including the Lamé constants, elastic modulus, Poisson's ratio, shear modulus, and bulk modulus. These parameters are frequently used to characterize the mechanical behavior of rocks and contribute significantly to brittleness assessments [12, 26-28].

Several brittleness indices have been developed UCS and indirect tensile strength measurements, such as those obtained from the Brazilian test [4, 15, 19] have proven highly useful for evaluating mechanical excavation performance. Brittleness indices, as a function of rock strength, are widely employed for the indirect estimation and evaluation of brittleness [29]. A review of the literature indicates that brittleness determination methods based on rock strength yield the most reliable results.

One of the most frequently used indicators of rock brittleness is the ratio of uniaxial compressive strength to tensile strength (σ_c / σ_t), commonly referred to as the brittleness ratio. Higher values of this ratio generally signify a more brittle material. Numerous researchers have proposed a variety of equations and definitions to quantify brittleness, several of which are summarized in Tab. 1 [30].

Table 1 Some empirical equations proposed for determining rock brittleness (modified from Meng et al. 2021) [24]

Proposed equation for the brittleness	Definition*	Parameters	Application areas**			Reference
			A	B	C	
$B_1 = \sigma_c / \sigma_t$	σ_c: Uniaxial compressive strength of the rock σ_t: tensile strength, respectively, r: Density of the rock sample, g: Unit weight or gravitational force acting on the rock	Strength parameters	√	√		[15]
$B_2 = (\sigma_c - \sigma_t) / (\sigma_c + \sigma_t)$			√	√		[15]
$B_3 = (\sigma_c + \sigma_t) / 2$				√		[31]
$B_4 = (\sigma_c . \sigma_t) / 2$				√		[5]
$B_5 = \sqrt{(\sigma_c . \sigma_t) / 2}$				√		[32]
$B_6 = 0.198\sigma_c - 2.174\sigma_t + 0.913p - 3.807$				√		[6]
$B_7 = 0.61\sigma_c^{0.81} - 1.37\sigma_t^{1.13} + 0.009\gamma^{2.72} - 5.45$				√		[33]
$B_8 = 0.59\sigma_c^{0.769} - 5.085\sigma_t^{0.531} + 0.09\gamma^{2.332}$				√		[34]
$B_9 = \left\{\ln\left[\tan(10.9+\gamma) + 53.27/\sigma_c + \tan\sigma_c + \sigma_c + 6.65\right]\right\}^2 + \left[(6.65 - \sigma_t + \tan\sigma_c + \sigma_c) / \left(4.17\sigma_t \sqrt[3]{\gamma^2/(\sigma_c - 4.17)}\right)\right]^2$				√		[34]
$B_{10} = (\sigma_t^{0.84} + E^{0.51}) / \sigma_t^{0.21}$	E: Elastic modulus, indicating the rock's stiffness			√		[35]
$B_{11} = 0.002\sigma_c / SIPS$	$SIPS$: the surface instability peak stress		√			[36]
$B_{12} = D\left(\sigma_t / K_{IC}\right)^2$	D: Structural dimension, such as beam height in fracture tests, K_{IC}: Fracture toughness, representing resistance to crack propagation					[36]
$B_{13} = K_b . K_s . P / h^2 = \sigma_t - K_b . P$	K_b: Brittleness index, used to quantify fracture tendency, K_s: Shape factor affecting load distribution, P: the applied load at failure, h: the distance between loading points, σ_t: the tensile strength	Point load test				[37]

*These parameters follow standard definitions used in rock mechanics literature but have been paraphrased for clarity.
**Application areas: (A) assessing the mechanical performance and structural stability of deep underground tunnels and mining environments, (B) evaluating rock fragmentation efficiency, which covers parameters such as drillability, cuttability, and penetration rate - all critical for mechanical excavation planning, and (C) analyzing the fracture potential (fracability) of unconventional hydrocarbon reservoirs during stimulation processes.It's important to note that indices not marked with a "√" symbol in the reference tables lack clearly defined application areas in their original sources.

In the literature, the brittleness of rocks has been estimated by some researchers using artificial neural networks. Shi et al. [38] attempted to predict the brittleness index, a critical parameter in characterizing shale gas reservoirs, using artificial neural networks and machine learning algorithms. In their study, Sun et al. [39] introduced a machine learning-based approach that employed several established algorithms - including the Chi-Square Automatic Interaction Detector (CHAID), Random Forest (RF), Support Vector Machine (SVM), K-Nearest Neighbors (KNN), and Artificial Neural Network (ANN) - to predict the brittleness index of rock samples and conducted a comparative analysis of their predictive performance. Hassan et al. [40] produced an artificial

intelligence-based model to determine rock brittleness based on the chemical properties of rocks. Lee and Lumley [41] utilized deep learning techniques to predict the mineralogical brittleness index of shales using seismic properties, Young's modulus, and Poisson's ratio. Zamanzadeh et al. [42] developed models to estimate the mechanical brittleness index by analyzing data from wells in oil fields, employing Least Squares Support Vector Machine (LSSVM), Multilayer Perceptron (MLP) neural networks, and hybrid forms of these with optimization algorithms such as Cuckoo Optimization Algorithm (COA), Particle Swarm Optimization (PSO), and Genetic Algorithm (GA). Xie et al. [43] explored effective artificial intelligence approaches to predict the brittleness and dynamic properties of limestones. Their study particularly focused on estimating dynamic properties such as the brittleness index, acoustic velocity, and elastic modulus.

Determining the physical and mechanical properties of rocks is crucial for engineering studies that interact with rock formations, such as underground and surface mining operations, underground openings, tunnels, dams, and foundations constructed within rock masses. In engineering applications, rock properties are primarily obtained through experimental and observational methods. Experimental methods involve tests conducted in accordance with rock mechanics standards, encompassing index and engineering properties that aid in the identification and correlation of rock masses and soils. However, obtaining, preparing, and conducting tests on rock samples can be costly and time-consuming for certain experiments [44]. To address these challenges, many researchers have adopted a range of indirect testing techniques to assess the engineering characteristics of rocks. These methods are generally straightforward, economical, and capable of delivering rapid results, making them well-suited for practical applications.

In the literature, no studies have been identified that specifically focus on predicting rock brittleness using artificial neural network-based programs and point load strength index values obtained from index tests. Within the scope of this study, point load strength index values derived from both the literature and laboratory experiments were used to predict the B_3 brittleness value proposed by [4] through machine learning techniques.

3 METHOD

The purpose of this study is to investigate and contrast how well various regression models predict B_3 values from PLI data. We aim to determine the best method for this predictive task by utilising a wide range of regression techniques, from sophisticated machine learning algorithms to conventional linear models.

3.1 Dataset

This study compiles data from a total of 520 rocks, including 287 sedimentary, 184 igneous, and 49 metamorphic rocks, sourced from the literature (Tab. 2).

The compiled dataset encompasses a diverse range of rock types, ensuring comprehensive representation in the analysis. This diversity provides a robust foundation for evaluating the relationship between Uniaxial Compressive Strength (UCS) and Point Load Index (PLI) values across different rock origins. Instead of a fixed split into training, validation, and test sets, cross-validation was employed to ensure robust model training and evaluation, as detailed in the "Model Training" section.

Table 2 Sources of data used in the study

Reference	Rock origin	Number of rocks
[44]	Igneous	3
	Metamorphic	1
	Sedimentary	3
[45]	Igneous	1
	Metamorphic	3
	Sedimentary	3
[46]	Sedimentary	4
[47]	Sedimentary	2
[48]	Sedimentary	6
[49]	Sedimentary	3
[50]	Igneous	4
	Metamorphic	1
	Sedimentary	7
[51]	Igneous	15
	Sedimentary	26
[52]	Igneous	1
[53]	Sedimentary	9
[54]	Sedimentary	6
[55]	Sedimentary	45
[56]	Igneous	2
[57]	Sedimentary	15
[58]	Sedimentary	10
[59]	Sedimentary	15
[60]	Igneous	32
[61]	Sedimentary	1
[62]	Sedimentary	23
[63]	Igneous	10
	Metamorphic	2
	Sedimentary	7
[64]	Sedimentary	1
[65]	Metamorphic	2
[66]	Sedimentary	1
[67]	Sedimentary	1
[68]	Igneous	19
	Metamorphic	20
	Sedimentary	17
[69]	Igneous	7
	Sedimentary	1
[70]	Sedimentary	30
[71]	Igneous	52
	Metamorphic	13
	Sedimentary	26
[72]	Metamorphic	7
	Sedimentary	3
[73]	Igneous	19
[74]	Igneous	18
	Sedimentary	11
[75]	Sedimentary	11
[76]	Igneous	1

The dataset supporting the findings of this study is not hosted as a single downloadable file. Instead, the original sources for all records are listed in Tab. 2; our analysis set comprises n = 520 curated entries obtained from an initial compilation of 523 records by removing one missing-PLI entry and pruning one exact duplicate measurement group. No winsorization or manual trimming was applied. The full preprocessing and modelling code is openly available at Kaggle (https://doi.org/10.34740/kaggle/ds/8292070).

It was analyzed 520 rock samples with the variables Reference (string identifier), Rock Origin (categorical: Sedimentary/Igneous/Metamorphic), UCS, BTS, PLI, and brittleness indices B_1-B_4 (numeric). One technical column

was empty and removed. Rows with missing values were excluded (one row with missing PLI), and two numerically duplicated records were pruned, yielding $n = 520$.

By rock type, the sample counts are Sedimentary 286, Igneous 191, Metamorphic 43. For the main variables used in the study, the distributions are summarized as follows:

- PLI ($n = 520$): median 4.14, IQR 2.42-6.69, min-max 0.07-15.73 (right-skewed).
- B_3 ($n = 520$): median 281.40, IQR 104.72-703.50, min-max 0.08-3388.40 (right-skewed).
- UCS ($n = 520$): median 72.05, IQR 39.20-122.34; BTS ($n = 520$): median 7.48, IQR 5.09-12.11.

To aid interpretation, we report per-group ranges for B_3 in the Results: Igneous 3.41-3388.40, Metamorphic 41.50-2621.87, Sedimentary 0.08-1810.00. A compact table of descriptive statistics (count, mean, SD, min, quartiles, max) can be seen in Tab. 3, and histograms of PLI and B_3 are provided in Fig. 1 and Fig. 2; the brittleness index by lithology (box plot) is shown in Fig. 3.

Table 3 Descriptive statistics (overall)

	Count	Mean	Std	Min	Q_1	Median	Q_3	Max
PLI	520.0	4.7774	2.9734	0.07	2.4175	4.14	6.69	15.73
B_3	520.0	498.29	558.31	0.08	104.72	281.40	703.50	3388.40
UCS	520.0	84.6456	52.7771	2.7	39.195	72.05	122.335	303.67
BTS	520.0	9.1145	5.4503	0.06	5.09	7.475	12.1125	34.4

As for data curation,we removed entries with missing PLI, pruned numerically duplicated records (identical UCS/BTS/PLI/BI), and standardized lithology labels. We did not winsorize or discard statistically extreme values; all physically plausible measurements were retained.

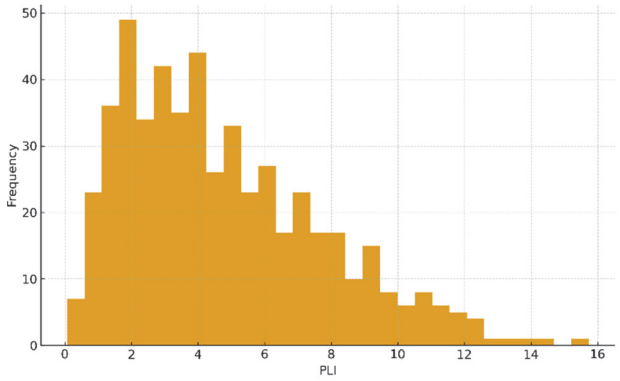

Figure 1 Distribution of PLI

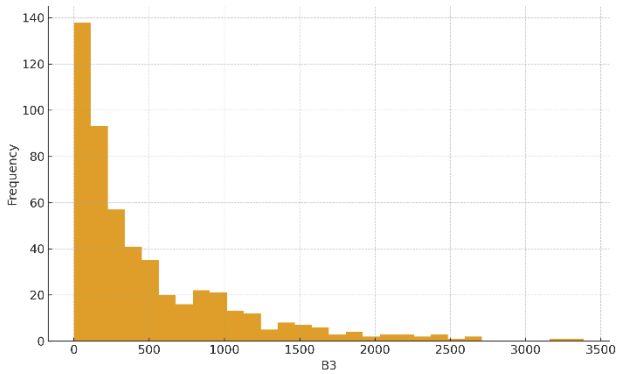

Figure 2 Distribution of B_3

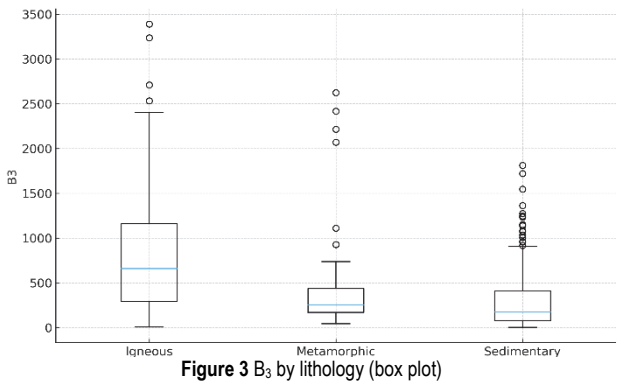

Figure 3 B_3 by lithology (box plot)

3.2 Model Selection

We evaluate a representative set of regression families linear and regularized linear (Linear, Ridge, Lasso, Elastic Net, Bayesian Ridge), tree-based/ensembles (Decision Tree, Random Forest, Extra Trees, Gradient Boosting, AdaBoost), kernel - and instance-based (SVR, KNN), a neural network (MLP), and a non-parametric statistical model (Gaussian Process) - to provide a balanced comparison when predicting B_3 from PLI alone. This portfolio spans interpretable linear baselines and non-linear learners that can model potentially complex PLI→B_3 relationships, includes methods with built-in variance/bias control and robustness to outliers, and follows a uniform evaluation protocol (5-fold cross-validation with grid-searched hyperparameters).

The 15 algorithms used to map PLI to B_3 are presented in Tab. 4. Linear Regression provides a baseline linear/monotonic fit; Ridge Regression applies L_2 shrinkage to reduce variance; Lasso Regression uses L_1 shrinkage as an alternative penalty; Elastic Net combines L_1 and L_2 penalties; Bayesian Ridge places Gaussian priors on coefficients to quantify parameter uncertainty; Huber Regressor uses the Huber loss for robustness to outliers; Support Vector Regression (SVR) captures smooth nonlinear trends via kernels and an ϵ-insensitive loss; k-Nearest Neighbors (KNN) performs local averaging along the PLI axis; Decision Tree partitions PLI by thresholds to produce piecewise-constant predictions; Random Forest averages bootstrapped trees to reduce variance; Extra Trees increase tree diversity through highly randomized splits; Gradient Boosting adds shallow trees sequentially to model residual structure; AdaBoost.R2 reweights samples to focus on hard cases in a boosting ensemble; Gaussian Process Regression (GPR) provides flexible kernel regression with predictive uncertainty; and Multilayer Perceptron (MLP) learns a flexible nonlinear mapping with appropriate regularization. In addition, the mathematical formulations corresponding to each regression model are also presented in Tab. 4.

Fig. 4 shows the architecture of Gradient Boosting Decision Tree (GBDT), highlighting its iterative nature. In this diagram, the process begins with an initial prediction, followed by a sequence of decision trees that progressively refine the model's performance by learning from the residual errors. Each tree contributes incrementally, with

the final model being the weighted sum of all individual trees. This visualization offers a clear understanding of how GBDT aggregates the strengths of multiple models to achieve robust predictive performance.

Table 4 The regression algorithms employed in this study and their mathematical formulations

Algorithm	Family	Equation
Linear Regression [77]	Linear	$\hat{y}_i = w_0 + w_1 x_i,\ \min_{w0,w1} = \sum_{i=1}^{n} \left(y_i - \left(w_0 + w_1 x_i\right)\right)^2$
Ridge Regression [78]	Linear + L_2	$\min_{w0,w1} = \sum_{i=1}^{n} \left(y_i - \left(w_0 + w_1 x_i\right)\right)^2 + \lambda w_{12}^2$
Lasso Regression [79]	Linear + L_1	$\min_{w0,w1} = \sum_{i=1}^{n} \left(y_i - \left(w_0 + w_1 x_i\right)\right)^2 + \lambda w_{11}$
Elastic Net [80]	Linear + L_1/L_2	$\min_{w0,w1} = \sum_{i=1}^{n} \left(y_i - \left(w_0 + w_1 x_i\right)\right)^2 + \lambda \left(\frac{1-\alpha}{2} w_{12}^2 + \alpha w_{11}\right)$
Bayesian Ridge [81]	Bayesian linear	$y \mid w, \sigma^2 \sim N\left(Xw, \sigma^2 I\right), w \sim N\left(0, \tau^{-1} I\right)$
Huber Regressor [82]	Robust linear	$\min_{w_0 w_1} = \sum_{i=1}^{n} \rho\delta\left(r_i\right), \rho\delta\left(r\right) = \begin{cases} \frac{1}{2} r^2, \lvert r\rvert \le \delta \\ \delta\left(\lvert r\rvert - \frac{1}{2}\delta\right), \lvert r\rvert \le \delta. \end{cases}$
SVR [83]	Kernel / margin	$\min_{w,b,\xi,\xi'} \frac{1}{2} w^2 + C \sum_i \left(\xi_i + \xi_i'\right)$
KNN [84]	Instance-based	$\hat{y}\left(x\right) = \frac{1}{K} \sum_{i \in N_K(x)} y_i$
Decision Tree [85]	Tree	$\hat{y}\left(x\right) = \sum_m \bar{y}_m 1\{x \in R_m\}$
Random Forest [86]	Ensemble (bagging)	$\hat{y}\left(x\right) = \frac{1}{B} \sum_{b=1}^{B} T_b\left(x\right)$
Extra Trees [87]	Ensemble (randomized)	Same as Random Forest's equation but thresholds/splits drawn at random before choosing the best among random candidates; prediction is the ensemble average.
Gradient Boosting [88]	Ensemble (boosting)	$f_0\left(x\right) = \text{argmin}_c \sum_i L\left(y_i, c\right)$
AdaBoost.R2 [89]	Ensemble (boosting)	$f_m\left(x\right) = \sum_{m=1}^{M} \alpha_m h_m\left(x\right), \alpha_m \alpha \ln \frac{1}{\beta_m}$
Gaussian Process [90]	Non-parametric Bayesian	$\hat{y}\left(x_*\right) = k_*\left(K + \sigma_n^2 I\right)^{-1} y,\ \text{Var}\left[\hat{y}\left(x_*\right)\right] = k\left(x_*, x_*\right) - k_*\left(K + \sigma_n^2 I\right)^{-1} k_*$
Multilayer Perceptron [91]	Neural network	$\hat{y} = W_L \sigma\left(W_{L-1} \sigma\left(\ldots \sigma\left(W_1 x + b_1\right)\ldots\right) + b_{L-1}\right) + b_L$

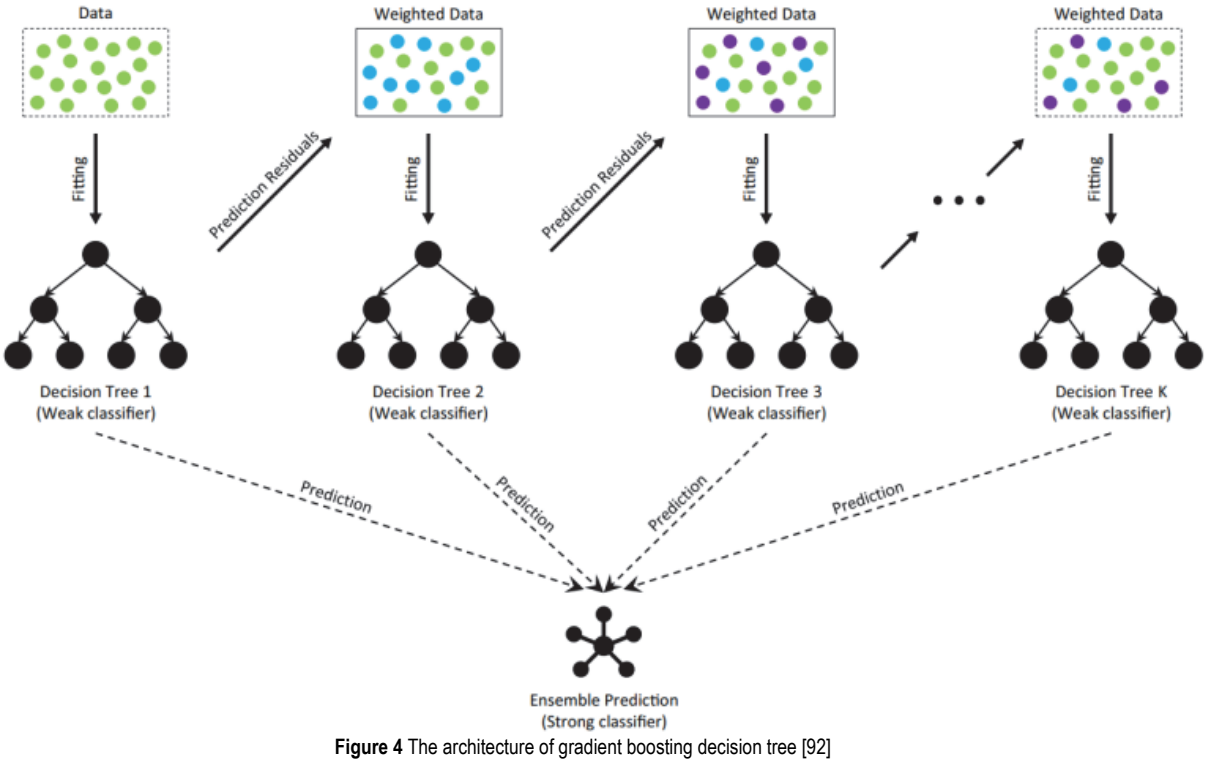

Figure 4 The architecture of gradient boosting decision tree [92]

3.3 Performance Metrics

To evaluate the performance and predictive reliability of the models, two key metrics were used: Root Mean Squared Error (*RMSE*) and the coefficient of determination (R^2). *RMSE* is a widely adopted measure for assessing the accuracy of predictions involving continuous variables. It is computed by taking the square root of the mean of the squared differences between the predicted and actual values. This metric reflects the magnitude of prediction errors, where smaller *RMSE* values indicate better model performance [93]. The formula for *RMSE* is presented in Eq. (1) [94].

$$RMSE = \sqrt{\frac{1}{m}\sum_{i=1}^{m}\left(X_i - Y_i\right)^2} \tag{1}$$

The coefficient of determination, denoted as R^2, quantifies how well the independent variables explain the variation observed in the dependent variable. It serves as a measure of the model's goodness of fit, where an R^2 value approaching 1.0 indicates that a large portion of the variance in the outcome variable is captured by the model's predictions [95]. The formula for *RMSE* is presented in Eq. (2) [94].

$$R^2 = 1 - \frac{\sum_{i=1}^{m}\left(X_i - Y_i\right)^2}{\sum_{i=1}^{m}\left(\overline{Y} - Y_i\right)^2} \tag{2}$$

RMSE and R^2 metrics together offer a comprehensive assessment of model performance, balancing both the accuracy of predictions and the explanatory power of the model. By employing *RMSE* and R^2, this study aims to provide a robust comparison of the different regression techniques, highlighting their relative strengths and weaknesses in predicting B_3 values based on PLI data.

3.4 Model Training

To enhance the reliability and generalizability of the model, a 5-fold cross-validation approach was applied. In this method, the dataset is partitioned into five equally sized subsets. During each iteration, the model is trained using four of these subsets, while the remaining one is used for validation. This procedure is repeated five times so that each subset is used once as the validation set. Furthermore, to optimize the hyperparameters of the models, we utilized Grid Search Cross-Validation (Grid Search CV). This method involves exhaustively searching over a specified parameter grid to find the best combination of hyperparameters that result in the highest model performance. Grid Search CV ensures that the selected hyperparameters are optimal for the given dataset and model [96].

Tab. 5 summarizes the hyperparameter search spaces used for model selection across all fifteen algorithms. For each method, we list the parameters tuned via grid search with 5-fold cross-validation and any fixed settings; empty grids indicate that no hyperparameters were tuned (e.g., Linear Regression, Bayesian Ridge). Identical CV splits were used for every model. For algorithms that require feature scaling (SVR, KNN, MLP), PLI was standardized within training folds only. Random seeds were fixed for tree/ensemble methods to ensure reproducibility. The grids exactly match those implemented in the accompanying notebooks; no settings beyond those listed were used.The Gradient Boosting algorithm yielded the best results in our experiments. The optimal hyperparameters identified through Grid Search CV were num_trees = 230 and learning_rate = 0.01. These settings enabled the model to achieve high accuracy while maintaining generalizability. To further illustrate the model's structure, a visualization of the first three trees in the ensemble is presented in Fig. 5, providing insights into the decision-making process and feature interactions.

Table 5 Hyperparameter grids

Algorithm	Hyperparameters
Linear Regression	{}
Random Forest Regressor	n_estimators: 8-24 (step 1); max_depth: [None, 2, 3, 4, 5, 6, 7, 8]
Gradient Boosting Regressor	n_estimators: 150-240 (step 10); learning_rate: [0.01, 0.02, 0.03]; max_depth: [3, 5, 7]
SVR	kernel: [linear, rbf]; C: [0.1, 0.2, 0.3, 5, 8, 10, 12]; gamma: [scale, auto]
Ridge	alpha: [0.1, 0.2, 0.3, 0.5, 5, 10]
Lasso	alpha: [0.1, 0.2, 0.3, 0.5, 5, 10]
Decision Tree Regressor	max_depth: [None, 10, 20, 30]; min_samples_split: [1, 2, 3, 5, 10, 20]; min_samples_leaf: [1, 3, 5, 7, 10]
KNN Regressor	n_neighbors: [1, 3, 5, 7, 9]; weights: [uniform, distance]
Gaussian Process Regressor	alpha: [$1e^{-2}$, $1e^{-3}$, $1e^{-4}$]
Elastic Net	alpha: [0.1-0.9 step 0.1]; l1_ratio: [0.1-0.9 step 0.1]
Bayesian Ridge	{}
Huber Regressor	alpha: [0.0001, 0.0002, 0.0003]
AdaBoost	n_estimators: 10-58 (step 3); learning_rate: [0.01, 0.02, 0.3]
Extra Trees	n_estimators: 150-290 (step 10); max_depth: [None, 5, 6, 7, 8, 10, 12, 20]
MLP Regressor	hidden_layer_sizes: [(100, 50), (75, 75), (50, 50), (75, 50), (100,), (20,10)]; activation: [relu, tanh]; solver: [adam, sgd]; alpha: [$1e^{-05}$, $2e^{-05}$]; learning_rate: [constant, adaptive]; learning_rate_init: [0.0001, 0.0002]; max_iter = 20000

3.5 Cross-Validation Protocol

We evaluate all models on the curated set of $n = 520$ pairs (x = PLI, $y = B_3$) using 5-fold cross-validation with a fixed random seed and identical folds across models: the dataset is randomly partitioned into five disjoint folds; for each fold j, the model is trained on the other four folds and predicted on fold j, with hyperparameters (if any) selected by grid search using only the training folds; we record *MSE*, *RMSE*, and R^2 on the held-out fold and repeat for all $j = 1, \ldots, 5$; finally, we aggregate the per-fold scores as mean ±SD (and, where needed, concatenate the out-of-fold predictions for diagnostic plots), and we rank models by the lowest mean CV-MSE to determine the best performer for the PLI→B_3 task.

An external hold-out set was not used because the available dataset is limited, and assembling additional, comparable measurements would be prohibitively costly and time-consuming.

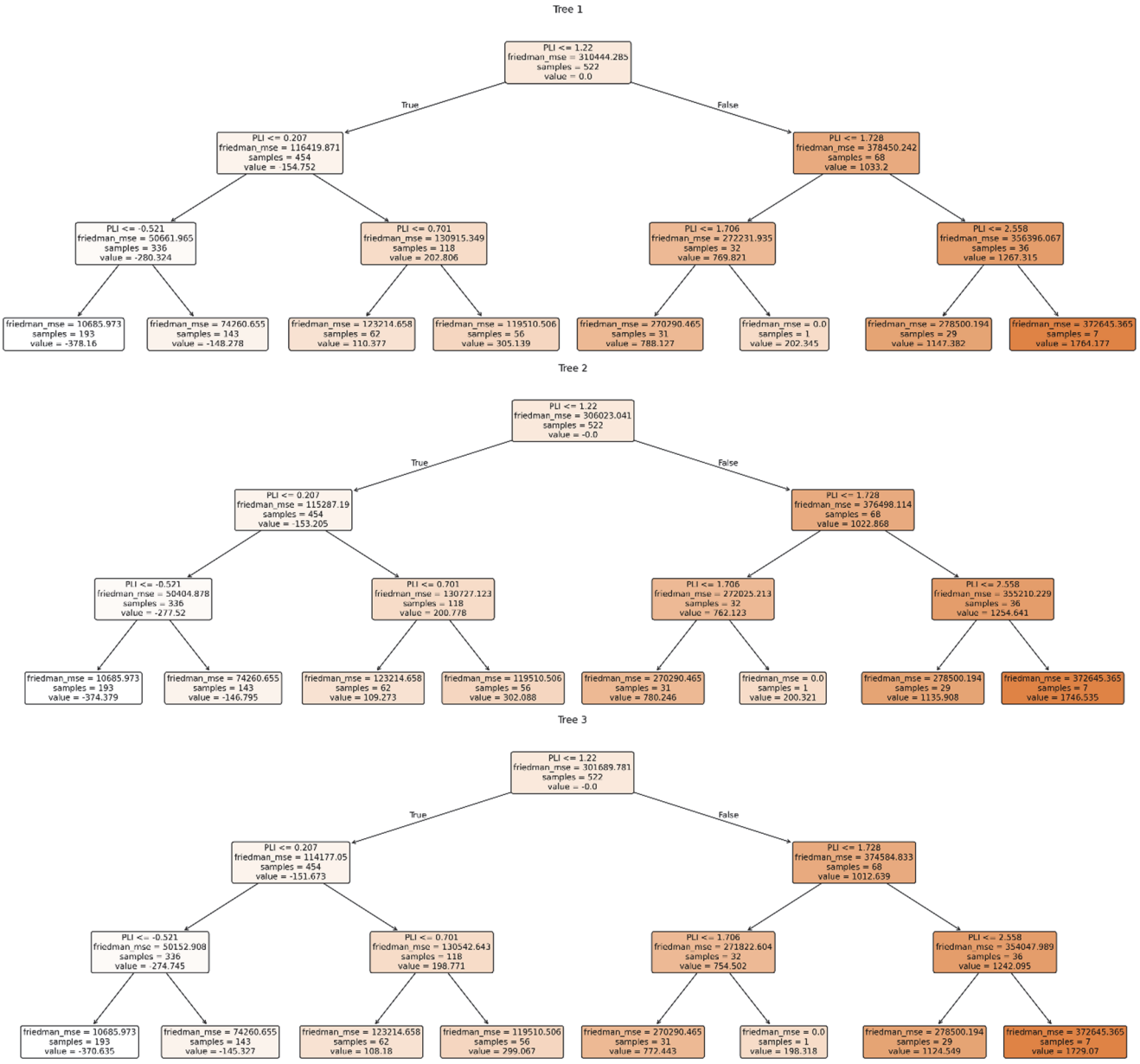

Figure 5 Visualization of the first three trees in the model

4 RESULTS

Although the model was trained using 15 different regressor models, namely Multilayer Perceptron, Linear Regression, Random Forest Regressor, Gradient Boosting, Support Vector Regressor, Ridge Regression, Lasso Regression, Decision Tree Regressor, K-Nearest Neighbors (KNN), Gaussian Process, Elastic Net, Bayesian Ridge, Huber Regressor, AdaBoost Regressor, and Extra Trees Regressor, the best results were achieved with Gradient Boosting. The results of the comparative analysis, conducted with cross-validation, are summarized in Tab. 5. For all types of rocks, under identical 5-fold CV on the PLI→B_3 task (see Tab. 6), Gradient Boosting ranked #1 across all lithologies (Sedimentary $R^2 = 0.81$, Igneous 0.72, Metamorphic 0.96) with the lowest *RMSE*s, consistent with threshold-like, piecewise trends from a single predictor. Random Forest was the overall #2 (Sedimentary $R^2 = 0.72$; Igneous 0.71; Metamorphic0.92), benefiting from variance reduction via tree averaging. The third-best varied by lithology: Linear Regression in Sedimentary ($R^2 = 0.65$) and Igneous (0.59), reflecting an approximately monotonic relation, whereas Gaussian Process in Metamorphic ($R^2 = 0.79$) slightly ahead of AdaBoost on *RMSE* - captured smooth nonlinearities. Exact metrics are reported in Tab. 6. To contextualize *RMSE*/*MSE* magnitudes, we provide the brittleness index distribution (Fig. 2) and per-lithology spread (Fig. 3), alongside descriptive statistics (Tab. 3).

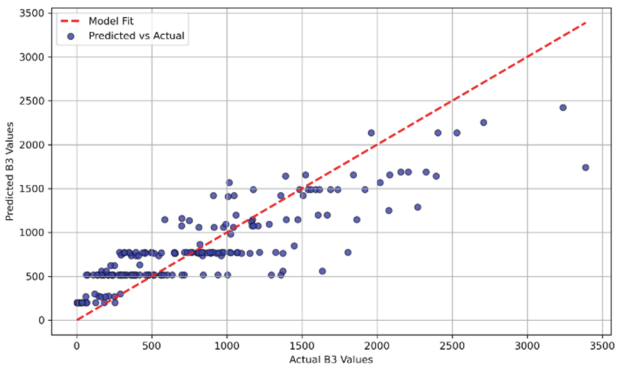

Figure 6 Predicted vs actual values for igneous

Fig. 6 displays a comparison between the predicted and actual brittleness values for igneous (magmatic) rock samples. Similarly, Fig. 7 illustrates this comparison for metamorphic rocks, while Fig. 8 provides the corresponding results for sedimentary rocks.

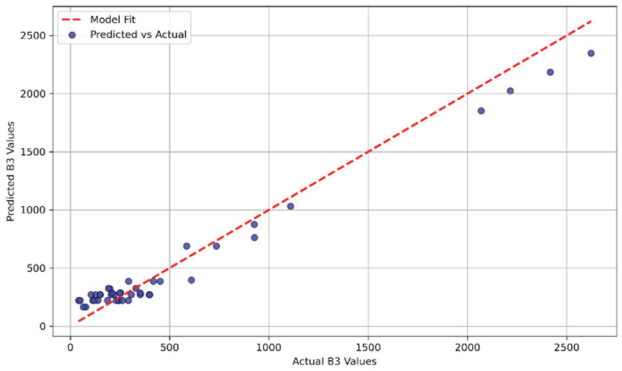

Figure 7 Predicted vs actual values for metamorphic

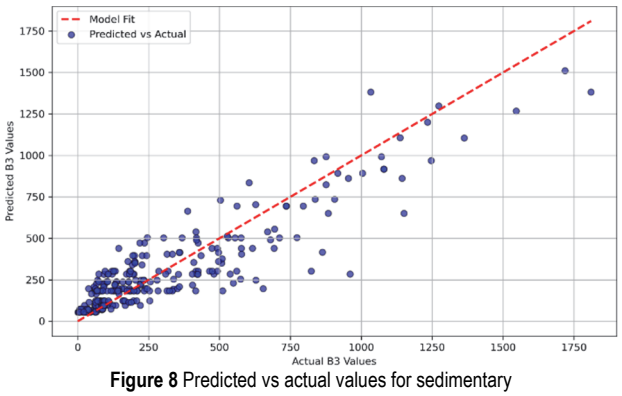

Figure 8 Predicted vs actual values for sedimentary

Tab. 6 summarizes the performance metrics for the Gradient Boosting algorithm, evaluated using 5-fold Cross-Validation on the entire dataset without considering the stone type. The Root Mean Square Error (*RMSE*) is 0.74, indicating that the model's predictions are, on average, close to the actual values, with low variability.

Table 6 Results with 5-fold cross-validation

	Sedimentary				Igneous				Metamorphic			
Algorithm	*RMSE*	*MSE*	*MAE*	R^2	*RMSE*	*MSE*	*MAE*	R^2	*RMSE*	*MSE*	*MAE*	R^2
Multilayer Perceptron	180.22	32479.25	134.28	0.61	422.65	178636.98	300.02	0.57	289.93	84060.19	197.50	0.38
Linear Regression	189.14	35773.94	135.01	0.65	421.31	177502.12	291.94	0.59	306.11	93703.33	242.35	0.77
Random Forest Regressor	168.92	28533.97	119.19	0.72	356.03	126757.36	256.82	0.71	178.19	31751.68	126.37	0.92
Gradient Boosting	140.05	19614	99.87	0.81	352.75	124432.56	267.34	0.72	119.11	14187.19	100.11	0.96
Support Vector Regressor	199.83	39932.03	132.56	0.61	440.16	193740.83	292.98	0.56	493.96	243996.48	252.80	0.39
Ridge Regression	196.17	38482.67	139.18	0.62	439.2	192896.64	308.94	0.56	392.12	153758.09	283.84	0.62
Lasso Regression	196.18	38486.59	139.18	0.62	440.32	193881.7	310.41	0.56	391.95	153624.8	284.09	0.62
Decision Tree Regressor	215.12	46276.61	148.38	0.55	448.09	200784.65	316.98	0.55	317.08	100539.73	220.61	0.75
KNN	211.06	44546.32	147.51	0.57	455.82	207771.87	320.93	0.53	324.95	105592.5	214.68	0.74
Gaussian Process	207.85	43201.62	143.13	0.58	435.59	189738.65	305.90	0.57	289.67	83908.71	177.70	0.79
Elastic Net	196.3	38533.69	138.99	0.62	439.19	192887.86	309.00	0.56	392.71	154221.14	283.22	0.62
Bayesian Ridge	196.22	38502.29	139.14	0.62	440.82	194322.27	311.09	0.56	392.26	153867.91	283.29	0.61
Huber Regressor	201.42	40570.02	139.11	0.6	437.02	190986.48	302.10	0.57	442.33	195655.83	286.99	0.51
AdaBoost Regressor	204.45	41799.8	142.58	0.59	437.81	191677.6	306.93	0.57	293.26	86001.43	176.15	0.79
ExtraTrees Regressor	194.21	37717.52	137.33	0.63	439.87	193485.62	309.55	0.56	316.7	100298.89	197.65	0.75

The Mean Squared Error (*MSE*) is 80,430.28, consistent with the data's scale. The Mean Absolute Error (*MAE*) of 188.85 reflects the average magnitude of prediction errors. The *R*-squared (R^2) value of 0.74 suggests that the model explains 74% of the variance in the data. These results, validated across five folds, confirm the robustness and reliability of the Gradient Boosting model in capturing the underlying patterns in the dataset, although some variance remains unexplained.

5 DISCUSSION AND CONCLUSION

Brittleness is a critical rock property that defines the susceptibility of rocks to fracture under applied loads, particularly in excavation operations. This characteristic is directly related to the durability of rocks. Consequently, predicting rock brittleness is essential for developing sustainable practices in numerous mining projects. In this study, brittleness values of rocks from various origins were predicted using point load strength index values and models derived from machine learning methods. Several models were developed, and efforts were made to identify the most suitable one.

Several prior studies have explored the application of machine learning techniques to predict rock brittleness using various input parameters. For instance, [38] and [39] employed artificial neural networks and ensemble learning methods to predict brittleness indices in shale gas

reservoirs, reporting promising accuracy levels. Similarly, [40] and [41] developed models based on elemental compositions and seismic-log data, respectively, further demonstrating the adaptability of AI models to diverse geological datasets. However, these studies generally rely on more complex or multi-parametric input features. In contrast, the present study emphasizes the effectiveness of using a single, easily obtainable index - Point Load Strength (PLI) - to achieve comparable or superior prediction performance, particularly through the Gradient Boosting algorithm. This streamlined approach offers practical advantages for field applications, especially when data availability is limited or rapid assessments are needed.

During the training of the model, we employed 5-fold cross-validation. To evaluate model performance, we used R^2, *MSE*, and *RMSE* metrics. R^2 is particularly crucial for assessing how well regression models explain the variance in the target variable. Using 5-fold cross-validation, the highest R^2 was achieved with Gradient Boosting due to its advanced ensemble learning capabilities, which effectively mitigate both bias and variance. Unlike the other 16 algorithms examined in this study, which often fail to capture complex, non-linear relationships inherent in the data, Gradient Boosting builds a series of decision trees sequentially, with each tree attempting to correct the errors of its predecessor. This iterative refinement allows the model to progressively enhance its predictive accuracy. Moreover, Gradient Boosting's inherent feature selection process identifies and leverages the most relevant features, further optimizing performance. Its robust handling of outliers and flexible hyperparameter tuning options enable it to achieve a balance between underfitting and overfitting, ensuring robust generalization on unseen data. Consequently, these attributes collectively contribute to the model's superior efficacy compared to other regression techniques evaluated in this study.

For metamorphic rocks, an R^2 of 0.96 indicates that the model explains 96% of the variance in B_3 values using only PLI values. For sedimentary rocks, this value reduces to 0.81, and for igneous rocks, it becomes 0.72. Although the acceptability of a high R^2 value depends on the context and field of study, generally a R^2 value greater than 0.7 is considered as acceptable. The *RMSE* and *MSE* results appear quite high; however, given the wide range of predicted values, this may be acceptable. The B_3 values range from 3.41 to 3388.4 for magmatic rocks, from 41.5 to 2621.87 for metamorphic rocks, and from 0.081 to 1810 for sedimentary rocks. Figs. 6, 7, and 8 present the predicted values from our model versus the actual values for each rock type. As shown, the model predicts B_3 values for metamorphic rocks with high accuracy using only PLI values. For sedimentary and magmatic rocks, while the accuracy decreases, it remains within an acceptable range.

While the proposed model demonstrates high predictive accuracy, especially for metamorphic rocks, it is important to acknowledge that its performance is inherently tied to the quality and consistency of the input PLI data. Point Load Strength Index (PLI) values can be affected by various factors such as specimen shape, size, surface conditions, and operator technique. The dataset used in this study aggregates values from multiple literature sources and experimental setups, which may introduce variability or measurement inconsistencies. Future studies could benefit from incorporating uncertainty quantification or sensitivity analyses to evaluate the robustness of predictions under different data quality scenarios. Standardizing PLI test procedures or using data augmentation techniques could also help mitigate this issue.

Another important consideration is that this study used only a single input variable. By focusing solely on the Point Load Strength Index (PLI), the model remains simple, interpretable, and suitable for practical field applications. However, this approach also limits the potential of more complex or hybrid machine learning and deep learning models, which typically require multiple features to fully leverage their predictive power. Future research could explore hybrid or deep learning architectures in multi-input scenarios to further improve prediction accuracy and handle more nuanced rock behaviour. Overall, while the simplicity of the model offers practical benefits, integrating additional features in future work could unlock further predictive improvements and allow for more advanced modelling techniques.

Prior work has used machine learning to predict brittleness with multi-parametric inputs drawn from shale-gas or limestone settings (e.g., elemental composition, seismic/elastic logs, conventional well logs), typically not PLI-only. In contrast, our study evaluates PLI→B_3 with a single predictor and reports lithology-aware performance (e.g., R^2 = 0.96 metamorphic; 0.81 sedimentary; 0.72 igneous). The Tab. 7 situates our results against representative studies mentioned in the manuscript. Across prior studies, both the target definition and the input modality vary substantially: Shi et al. [38] and Lee & Lumley [41] predict log-based brittleness from well/seismic logs; Hassan et al. [40] targets a mineralogical brittleness index from elemental/XRF inputs; Sun et al. [39] uses mechanical brittleness but with multi-parameter physical tests (e.g., Schmidt hammer, P-wave, density). By contrast, this study addresses a strictly mechanical brittleness (B_3) target using a single, field-practical input (PLI) and evaluates 15 representative regression families under identical 5-fold CV across three lithologies (Sedimentary/Igneous/Metamorphic). The curated dataset (n = 520, compiled from ≈40 sources) is the largest among the compared works and uniquely enables a single-predictor benchmark that is directly actionable on site. In short, while prior art spans heterogeneous targets and richer input suites, our contribution isolates the PLI→B_3 mapping at scale, clarifies which model families generalize best, and provides a lightweight pathway for engineering use.

For Practical application in operational settings, the workflow is: (1) perform the point-load test on intact specimens (site or laboratory) and compute PLI per the standard procedure; (2) feed the measured PLI to the selected regression model from this study to obtain the estimated brittleness index B_3; (3) interpret the estimate within the empirical range of our data (PLI ≈0.07–15.73) and avoid extrapolation outside this range; (4) where lithology is known (Sedimentary/Igneous/Metamorphic), use the corresponding performance summaries reported in this paper to contextualize expected uncertainty; (5) for critical decisions, corroborate with a small subset of confirmatory UCS/BTS tests when feasible. This

procedure preserves the practicality of PLI while providing an immediate, single-input brittleness estimate for engineering planning. As dataset-specific guidance from our learning-curve analysis (PLI→B_3, 5-fold CV), diminishing returns occur at roughly n≈200 overall; by lithology, ≈45 (Sedimentary), ≈30 (Igneous), and ≈20 (Metamorphic). These are practical targets rather than strict minima.

A key limitation is the lack of external validation, primarily due to the small size of the curated dataset and the substantial effort and cost required to obtain further comparable data; future work will evaluate the models on an independent set.

Table 7 Comparative summary of ML studies on brittleness prediction vs. presented study

Study	Algorithms	Target
Shi et al., 2016 [38]	BP-ANN, ELM, Linear Regression	Log-based brittleness index (from conventional well logs).
Sun et al., 2020 [39]	CHAID, Random Forest, SVM, KNN, ANN	Rock brittleness index (mechanical) from physical/indirect tests
Hassan et al., 2022 [40]	ANN, Fuzzy Logic, SVM (inputs: elemental/XRF)	Mineralogical brittleness index (MBI) from elemental composition (XRF).
Lee & Lumley, 2023 [41]	MLR, Decision Trees, Ensembles (bagged/boosted), SVM, Neural (incl. DFNN)	Log-based brittleness index (LBI) from seismic/elastic logs
Zamanzadeh T. et al., 2023 [42]	LSSVM, MLP, hybrids (COA/PSO/GA)	Well-log–based mechanical brittleness index
Xie et al., 2024 [43]	Simple/Multiple Linear Regression, GPR, MFFNN, SVM	Brittleness index and dynamic properties for limestones.
This Study	15 regressors across families (linear/regularized, robust, SVM, KNN, trees/ensembles, GP, MLP); best = Gradient Boosting	B_3 (mechanical brittleness index) from PLI (single input)

6 REFERENCES

[1] Zhang, D., Ranjith, P., & Perera, M. (2016). The brittleness indices used in rock mechanics and their application in shale hydraulic fracturing: A review. *Journal of Petroleum Science and Engineering*, *143*, 158-170. https://doi.org/10.1016/j.petrol.2016.02.011

[2] Xuefeng, L., Shibo, W., Shirong, G., Malekian, R., & Zhixiong, L. (2018). Investigation on the influence mechanism of rock brittleness on rock fragmentation and cutting performance by discrete element method. *Measurement*, *113*, 120-130. https://doi.org/10.1016/j.measurement.2017.07.043

[3] Blindheim, O. & Bruland, A. (1998). Boreability testing. *Norwegian TBM Tunneling 30 Years of Experience with TBMs in Norwegian Tunneling*, 21-27.

[4] Altindag, R. (2002). The evaluation of rock brittleness concept on rotary blast hold drills. *Journal of the Southern African Institute of Mining and Metallurgy*, *102*(1), 61-66.

[5] Altindag, R. (2003). Correlation of specific energy with rock brittleness concepts on rock cutting.*Journal of the Southern African Institute of Mining and Metallurgy*, *103*(3), 163-171.

[6] Yagiz, S. (2008). Utilizing rock mass properties for predicting TBM performance in hard rock condition. *Tunnelling and Underground Space Technology*, *23*(3), 326-339. https://doi.org/10.1016/j.tust.2007.04.011

[7] Liu, Q., Liu, J., Shi, K., Pan, Y., Huang, X., Liu, X., & Wei, L. (2016). Evaluation of rock brittleness indexes on rock fragmentation efficiency by disc cutter. *Chinese Journal of Rock Mechanics and Engineering*, *35*(3), 498-510.

[8] Jarvie, D. M., Hill, R. J., Ruble, T. E., & Pollastro, R. M. (2007). Unconventional shale-gas systems: The Mississippian Barnett Shale of north-central Texas as one model for thermogenic shale-gas assessment. *AAPG Bulletin*, *91*(4), 475-499. https://doi.org/10.1306/12190606068

[9] Rickman, R., Mullen, M., Petre, E., Grieser, B., & Kundert, D. (2008, September 21-24). *A practical use of shale petrophysics for stimulation design optimization: All shale plays are not clones of the Barnett Shale*. SPE Annual Technical Conference and Exhibition, Denver, Colorado, USA. https://doi.org/10.2118/115258-MS

[10] Jin, X., Shah, S., Truax, J., & Roegiers, J.-C. (2014, October 27-29). *A practical petrophysical approach for brittleness prediction from porosity and sonic logging in shale reservoirs*. SPE Annual Technical Conference and Exhibition, Amsterdam, Netherlands. https://doi.org/10.2118/170972-MS

[11] Jin, X., Shah, S. N., Roegiers, J.-C., & Zhang, B. (2014, February 4-6). *Fracability evaluation in shale reservoirs-an integrated petrophysics and geomechanics approach. SPE Hydraulic Fracturing* Technology Conference, Woodlands, Texas, USA. https://doi.org/10.2118/168589-MS

[12] Guo, T., Zhang, S., Ge, H., Wang, X., Lei, X., & Xiao, B. (2015). A new method for evaluation of fracture network formation capacity of rock. *Fuel*, *140*, 778-787. https://doi.org/10.1016/j.fuel.2014.10.017

[13] Kim, T., Hwang, S., & Jang, S. (2017). Petrophysical approach for S-wave velocity prediction based on brittleness index and total organic carbon of shale gas reservoir: A case study from Horn River Basin, Canada. *Journal of Applied Geophysics*, *136*, 513-520. https://doi.org/10.1016/j.jappgeo.2016.12.003

[14] Hajiabdolmajid, V. & Kaiser, P. (2003). Brittleness of rock and stability assessment in hard rock tunneling. *Tunnelling and Underground Space Technology, 18*(1), 35-48. https://doi.org/10.1016/S0886-7798(02)00100-1

[15] Hucka, V. & Das, B. (1974). Brittleness determination of rocks by different methods. *International Journal of Rock Mechanics and Mining Sciences & Geomechanics Abstracts*, *11*(10), 389-392. https://doi.org/10.1016/0148-9062(74)91109-7

[16] Singh, S. (1986). Brittleness and the mechanical winning of coal. *Mining Science and Technology*, *3*(3), 173-180. https://doi.org/10.1016/S0167-9031(86)90305-1

[17] Göktan, R. (1991). Brittleness and micro-scale rock cutting efficiency.*Mining Science and Technology*, *13*(3), 237-241. https://doi.org/10.1016/0167-9031(91)90339-E

[18] Aubertin, M., Gill, D. E., & Simon, R. (1994, June 1–3). *On the use of the brittleness index modified (BIM) to estimate the post-peak behavior of rocks*. 1st North American Rock Mechanics Symposium, Austin, Texas, USA.

[19] Yagiz, S. (2006, Septemper 6-10). *A model for prediction of tunnel boring machine performance. Substructures and underground space. Engineering geology for tomorrow's cities*. The 10th International Association of Engineering Geologists Congress, Nottingham, United Kingdom.

[20] Yarali, O. (2007, June 6-8). *Investigation of the relations between rock brittleness and drilling rate index*. The 20th International Mining Congress and Exhibition of Turkey, Ankara, Turkey.

[21] Gunes Yilmaz, N., Karaca, Z., Goktan, R. M., & Akal, C. (2009). Relative brittleness characterization of some selected

granitic building stones: influence of mineral grain size. *Construction and Building Materials*, *23*(1), 370-375. https://doi.org/10.1016/j.conbuildmat.2007.11.014

[22] Altındağ, R. & Güney, A. (2010). Predicting the relationships between brittleness and mechanical properties (UCS, TS and SH) of rocks. *Scientific Research and Essays*, *5*(16), 2107-2118.

[23] Yagiz, S. (2004, March 8-12). *Correlation between uniaxial compressive strength and brittleness of selected rock types*. 57th Geological Cogress of Turkey - Ankara, Ankara, Turkey.

[24] Meng, F., Wong, L. N. Y., & Zhou, H. (2021). Rock brittleness indices and their applications to different fields of rock engineering: A review. *Journal of Rock Mechanics and Geotechnical Engineering*, *13*(1), 221-247. https://doi.org/10.1016/j.jrmge.2020.06.008

[25] Andreev, G. E., (1995). *Brittle failure of rock materials*. CRC press.

[26] Huang, X.-R., Huang, J.-P., Li, Z.-C., Yang, Q.-Y., Sun, Q.-X., & Cui, W. (2015). Brittleness index and seismic rock physics model for anisotropic tight-oil sandstone reservoirs. *Applied Geophysics*, *12*(1), 11-22. https://doi.org/10.1007/s11770-014-0478-0

[27] Quinn, J. & Quinn, G. (1997). Indentation brittleness of ceramics: a fresh approach. *Journal of Materials Science*, *32*, 4331-4346. https://doi.org/10.1023/A:1018671823059

[28] Sone, H. & Zoback, M. D. (2013). Mechanical properties of shale-gas reservoir rocks-Part 1: Static and dynamic elastic properties and anisotropy. *Geophysics*, *78*(5), D381-D392. https://doi.org/10.1190/GEO2013-0050.1

[29] Kaunda, R. B. & Asbury, B. (2016). Prediction of rock brittleness using nondestructive methods for hard rock tunneling. *Journal of Rock Mechanics and Geotechnical Engineering*, *8*(4), 533-540. https://doi.org/10.1016/j.jrmge.2016.03.002

[30] Akbay, D., Ekincioğlu, G., Altındağ, R., & Şengün, N. (2021). Farklı cihaz ve yöntemler ile belirlenen Shore sertlik değerlerinin karbonatlı kayaçların gevreklik değerlerinin tahmininde kullanılabilirliğinin incelenmesi. *Pamukkale Üniversitesi Mühendislik Bilimleri Dergisi*, *27*(3), 441-448.

[31] Özfirat, M. K., Yenice, H., Şimşir, F., & Yaralı, O. (2016). A new approach to rock brittleness and its usability at prediction of drillability. *Journal of African Earth Sciences*, *119*, 94-101. https://doi.org/10.1016/j.jafrearsci.2016.03.017

[32] Altindag, R. (2010). Assessment of some brittleness indexes in rock-drilling efficiency. *Rock Mechanics and Rock Engineering*, *43*, 361-370. https://doi.org/10.1007/s00603-009-0057-x

[33] Yagiz, S. & Gokceoglu, C. (2010). Application of fuzzy inference system and nonlinear regression models for predicting rock brittleness. *Expert Systems with Applications*, *37*(3), 2265-2272. https://doi.org/10.1016/j.eswa.2009.07.046

[34] Khandelwal, M., Shirani Faradonbeh, R., Monjezi, M., Armaghani, D.J., Majid, M.Z.B.A., & Yagiz, S. (2017). Function development for appraising brittleness of intact rocks using genetic programming and non-linear multiple regression models. *Engineering with Computers*, *33*, 13-21. https://doi.org/10.1007/s00366-016-0452-3

[35] Nejati, H. & Moosavi, S.A. (2017). A new brittleness index for estimation of rock fracture toughness. *Journal of Mining and Environment*, *8*(1), 83-91.

[36] Tarokh, A., Peng, J., Fakhimi, A., & Labuz, J. (2016, June 26-29). *Evaluation of brittleness from spalling and bending tests*. 50th US Rock Mechanics / Geomechanics Symposium, Houston, Texas, USA.

[37] Reichmuth, D. R. (1967, April 17-19). *Point load testing of brittle materials to determine tensile strength and relative brittleness*. 9th U.S. Symposium on Rock Mechanics, Golden, Colorado, USA.

[38] Shi, X., Liu, G., Cheng, Y., Yang, L., Jiang, H., Chen, L., Jiang, S., & Wang, J. (2016). Brittleness index prediction in shale gas reservoirs based on efficient network models. *Journal of Natural Gas Science and Engineering*, *35*, 673-685. https://doi.org/10.1016/j.jngse.2016.09.009

[39] Sun, D., Lonbani, M., Askarian, B., Jahed Armaghani, D., Tarinejad, R., Thai Pham, B., & Huynh, V. V. (2020). Investigating the applications of machine learning techniques to predict the rock brittleness index. *Applied Sciences*, *10*(5), 1691. https://doi.org/10.3390/app10051691

[40] Hassan, A., Chan, S., Mahmoud, M., Aljawad, M.S., Humphrey, J., & Abdulraheem, A. (2022). Artificial intelligence-based model of mineralogical brittleness index based on rock elemental compositions. *Arabian Journal for Science and Engineering*, *47*(9), 11745-11761. https://doi.org/10.1007/s13369-021-06487-6

[41] Lee, J. & Lumley, D. E. (2023). Predicting shale mineralogical brittleness index from seismic and elastic property logs using interpretable deep learning. *Journal of Petroleum Science and Engineering*, *220*, 111231. https://doi.org/10.1016/j.petrol.2022.111231

[42] Zamanzadeh Talkhouncheh, M., Davoodi, S., Larki, B., Mehrad, M., Rashidi, S., & Vasfi, M. (2023). A new approach to mechanical brittleness index modeling based on conventional well logs using hybrid algorithms. *Earth Science Informatics*, *16*(4), 3387-3416. https://doi.org/10.1007/s12145-023-01098-1

[43] Xie, Y., Wang, L., Gu, Y., Gu, X., Chen, S., Khajehzadeh, M., & Hosseini, S. (2024). Prediction of Rock's Brittleness and Dynamic Properties Utilizing Effective Artificial Intelligence Approaches. *Periodica Polytechnica Civil Engineering*, *68*(3), 946-960. https://doi.org/10.3311/PPci.23156

[44] Akbay, D. (2018). *Nokta yükleme deneyinde yapılan hataların giderilmesi için yeni bir deney düzeneğinin tasarlanması*. Doctoral dissertation, Süleyman Demirel University.

[45] Altindag, R. (2000). The relationships between fracture toughness and other mechancical properties rocks. *Journal of Science Engineering*, *2*(2), 39-47.

[46] Aston, T., MacIntyre, J., & Kazi, H. (1991). The effect of worn and chipped points on point load indices. *Mining Science and Technology*, *13*(1), 69-74. https://doi.org/10.1016/0167-9031(91)90268-H

[47] Balcı, C. & Bilgin, N. (2010). Mekanize kazı makinalarının seçiminde küçük ve tam boyutlu kazı deneylerinin karşılaştırılması. *itüdergisi/d*, *4*(3), 76-86.

[48] Basarir, H. & Karpuz, C. (2004). A rippability classification system for marls in lignite mines. *Engineering Geology*, *74*(3-4), 303-318. https://doi.org/10.1016/j.enggeo.2004.04.004

[49] Basarir, H., Karpuz, C., & Bozdag, T., (2018). Rippability assessment studies at Tunçbilek Coal Mine: A case study. *Mine Planning and Equipment Selection 2000*, 515-520. https://doi.org/10.1201/9780203747124-98

[50] Bearman, R. (1999). The use of the point load test for the rapid estimation of Mode I fracture toughness. *International Journal of Rock Mechanics and Mining Sciences*, *36*(2), 257-263. https://doi.org/10.1016/S0148-9062(99)00015-7

[51] Capik, M. & Yilmaz, A. O. (2017). Modeling of Micro Deval abrasion loss based on some rock properties. *Journal of African Earth Sciences*, *134*, 549-556. https://doi.org/10.1016/j.jafrearsci.2017.04.006

[52] Fakir, M., Ferentinou, M., & Misra, S. (2017). An investigation into the rock properties influencing the strength in some granitoid rocks of KwaZulu-Natal, South Africa. *Geotechnical and Geological Engineering*, *35*, 1119-1140. https://doi.org/10.1007/s10706-017-0168-1

[53] Ghobadi, M. & Babazadeh, R. (2015). Experimental studies on the effects of cyclic freezing-thawing, salt crystallization, and thermal shock on the physical and mechanical

characteristics of selected sandstones. *Rock Mechanics and Rock Engineering*, *48*, 1001-1016. https://doi.org/10.1007/s00603-014-0609-6

[54] Gunsallus, K. L. & Kulhawy, F. H. (1984). A comparative evaluation of rock strength measures. *International Journal of Rock Mechanics and Mining Sciences & Geomechanics Abstracts*, *21*(5), 233-248. https://doi.org/10.1016/0148-9062(84)92680-9

[55] Heidari, M., Khanlari, G., Torabi Kaveh, M., & Kargarian, S. (2012). Predicting the uniaxial compressive and tensile strengths of gypsum rock by point load testing. *Rock Mechanics and Rock Engineering*, *45*, 265-273. https://doi.org/10.1007/s00603-011-0196-8

[56] Heidari, M., Momeni, A., & Naseri, F. (2013). New weathering classifications for granitic rocks based on geomechanical parameters. *Engineering Geology*, *166*, 65-73. https://doi.org/10.1016/j.enggeo.2013.08.007

[57] Jamshidi, A., Nikudel, M.R., Khamehchiyan, M., Zarei Sahamieh, R., & Abdi, Y. (2016). A correlation between P-wave velocity and Schmidt hardness with mechanical properties of travertine building stones. *Arabian Journal of Geosciences*, *9*, 1-12. https://doi.org/10.1007/s12517-016-2542-3

[58] Jamshidi, A., Abdi, Y., & Sarikhani, R. (2020). Prediction of brittleness indices of Sandstones using a Novel Physico-Mechanical parameter. *Geotechnical and Geological Engineering*, *38*, 4651-4659. https://doi.org/10.1007/s10706-020-01316-3

[59] Kahraman, S., Balcı, C., Yazıcı, S., & Bilgin, N. (2000). Prediction of the penetration rate of rotary blast hole drills using a new drillability index. *International Journal of Rock Mechanics and Mining Sciences*, *37*(5), 729-743. https://doi.org/10.1016/S1365-1609(00)00007-1

[60] Karaman, K., Kesimal, A., & Ersoy, H. (2015). A comparative assessment of indirect methods for estimating the uniaxial compressive and tensile strength of rocks. *Arabian Journal of Geosciences*, *8*, 2393-2403. https://doi.org/0.1007/s12517-014-1384-0

[61] Khajevand, R. & Fereidooni, D. (2018). Assessing the empirical correlations between engineering properties and P wave velocity of some sedimentary rock samples from Damghan, northern Iran. *Arabian Journal of Geosciences*, *11*, 1-17. https://doi.org/10.1007/s12517-018-3810-1

[62] Khanlari, G.-R., Heidari, M., Sepahigero, A.-A., & Fereidooni, D. (2014). Quantification of strength anisotropy of metamorphic rocks of the Hamedan province, Iran, as determined from cylindrical punch, point load and Brazilian tests. *Engineering Geology*, *169*, 80-90. https://doi.org/10.1016/j.enggeo.2013.11.014

[63] Kılıç, A. & Teymen, A. (2008). Determination of mechanical properties of rocks using simple methods. *Bulletin of Engineering Geology and the Environment*, *67*, 237-244. https://doi.org/10.1007/s10064-008-0128-3

[64] Lashkaripour, G. R. & Nakhaei, M. (2001). *A statistical investigation on mudrocks*. ISRM Regional Symposium EUROCK 2001, Espoo, Finland.

[65] Li, D. & Wong, L.N.Y. (2013). Point load test on meta-sedimentary rocks and correlation to UCS and BTS. *Rock Mechanics and Rock Engineering*, *46*, 889-896. https://doi.org/10.1007/s00603-012-0299-x

[66] Masoumi, H., Horne, J., & Timms, W. (2017). Establishing empirical relationships for the effects of water content on the mechanical behavior of Gosford sandstone. *Rock Mechanics and Rock Engineering*, *50*, 2235-2242. https://doi.org/10.1007/s00603-017-1243-x

[67] Minaeian, B. & Ahangari, K. (2017). Prediction of the uniaxial compressive strength and Brazilian tensile strength of weak conglomerate. *International Journal of Geo-Engineering*, *8*(1), 19. https://doi.org/10.1186/s40703-017-0056-9

[68] Mishra, D. A. & Basu, A. (2012). Use of the block punch test to predict the compressive and tensile strengths of rocks. *International Journal of Rock Mechanics and Mining Sciences*, *51*, 119-127. https://doi.org/10.1016/j.ijrmms.2012.01.016

[69] Singh, R., Umrao, R. K., Ahmad, M., Ansari, M., Sharma, L., & Singh, T. (2017). Prediction of geomechanical parameters using soft computing and multiple regression approach. *Measurement*, *99*, 108-119. https://doi.org/10.1016/j.measurement.2016.12.023

[70] Tahir, M., Mohammad, N., & Din, F. (2011). Strength parameters and their inter-relationship for limestone of Cherat and Kohat areas of Khyber Pakhtunkhwa. *Journal of Himalayan Earth Science*, *44*(2), 45-51.

[71] Teymen, A. & Mengüç, E. C. (2020). Comparative evaluation of different statistical tools for the prediction of uniaxial compressive strength of rocks. *International Journal of Mining Science and Technology*, *30*(6), 785-797. https://doi.org/10.1016/j.ijmst.2020.06.008

[72] Tripathy, A., Singh, T., & Kundu, J. (2015). Prediction of abrasiveness index of some Indian rocks using soft computing methods. *Measurement*, *68*, 302-309. https://doi.org/10.1016/j.measurement.2015.03.009

[73] Tuğrul, A. & Zarif, I. (1999). Correlation of mineralogical and textural characteristics with engineering properties of selected granitic rocks from Turkey. *Engineering Geology*, *51*(4), 303-317. https://doi.org/10.1016/S0013-7952(98)00071-4

[74] Yarali, O. & Soyer, E. (2013). Assessment of relationships between drilling rate index and mechanical properties of rocks. *Tunnelling and Underground Space Technology*, *33*, 46-53. https://doi.org/10.1016/j.tust.2012.08.010

[75] Yenice, H. (2002). Bazi Kayaçlarin Tek Eksenli Basinç Dayanimlari ile Diğer Malzeme Özellikleri Arasindaki İlişkiler. *Dokuz Eylül Üniversitesi Mühendislik Fakültesi Fen ve Mühendislik Dergisi*, *4*(2), 65-71.

[76] Yesiloglu-Gultekin, N., Gokceoglu, C., & Sezer, E.A. (2013). Prediction of uniaxial compressive strength of granitic rocks by various nonlinear tools and comparison of their performances. *International Journal of Rock Mechanics and Mining Sciences*, *62*, 113-122. https://doi.org/10.1016/j.ijrmms.2013.05.005

[77] Montgomery, D. C., Peck, E. A., & Vining, G. G., (2021). *Introduction to linear regression analysis*. John Wiley & Sons.

[78] Hoerl, A. E. & Kennard, R. W. (1970). Ridge regression: Biased estimation for nonorthogonal problems. *Technometrics*, *12*(1), 55-67. https://doi.org/10.1080/00401706.1970.10488634

[79] Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, *58*(1), 267-288. https://doi.org/10.1111/j.2517-6161.1996.tb02080.x

[80] Zou, H. & Hastie, T. (2005). Regularization and variable selection via the elastic net. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, *67*(2), 301-320. https://doi.org/10.1111/j.1467-9868.2005.00503.x

[81] MacKay, D. J. C. (1992). Bayesian interpolation. *Neural Computation*, *4*(3), 415-447. https://doi.org/10.1162/neco.1992.4.3.415

[82] Huber, P. J. (1992). Robust estimation of a location parameter. *Breakthroughs in Statistics: Methodology and Distribution*, 492-518. https://doi.org/10.1007/978-1-4612-4380-9_35

[83] Smola, A. J. & Schölkopf, B. (2004). A tutorial on support vector regression. *Statistics and Computing*, *14*, 199-222. https://doi.org/10.1023/B:STCO.0000035301.49549.88

[84] Cover, T. & Hart, P. (1967). Nearest neighbor pattern classification. *IEEE Transactions on Information Theory*, *13*(1), 21-27. https://doi.org/10.1109/TIT.1967.1053964

[85] Breiman, L., Friedman, J., Olshen, R. A., & Stone, C. J. (1984). *Classification and Regression Trees.* Chapman and Hall/CRC.
[86] Breiman, L. (2001). Random forests. *Machine Llearning*, *45*, 5-32. https://doi.org/10.1023/A:1010933404324
[87] Geurts, P., Ernst, D., & Wehenkel, L. (2006). Extremely randomized trees. *Machine Learning*, *63*, 3-42. https://doi.org/10.1007/s10994-006-6226-1
[88] Friedman, J. H. (2001). Greedy function approximation: A gradient boosting machine. *The Annals of Statistics*, *29*(5), 1189-1232. https://doi.org/10.1214/aos/1013203451
[89] Freund, Y. & Schapire, R. E. (1997). A decision-theoretic generalization of on-line learning and an application to boosting. *Journal of Computer and System Sciences*, *55*(1), 119-139. https://doi.org/10.1006/jcss.1997.1504
[90] Rasmussen, C. E. & Williams, C. K. I. (2006). *Gaussian processes for machine learning*. MIT Press. https://doi.org/10.7551/mitpress/3206.001.0001
[91] Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep Learning*. MIT Press.
[92] Deng, H., Zhou, Y., Wang, L., & Zhang, C. (2021). Ensemble learning for the early prediction of neonatal jaundice with genetic features. *BMC Medical Informatics and Decision Making*, *21*(1), 338. https://doi.org/10.1186/s12911-021-01701-9
[93] Chai, T. & Draxler, R. R. (2014). Root mean square error (RMSE) or mean absolute error (MAE). *Geoscientific Model Development Discussions*, *7*(1), 1525-1534. https://doi.org/10.5194/gmd-7-1247-2014
[94] Chicco, D., Warrens, M. J., & Jurman, G. (2021). The coefficient of determination R-squared is more informative than SMAPE, MAE, MAPE, MSE and RMSE in regression analysis evaluation. *PeerJ Computer Science*, *7*, e623. https://doi.org/10.7717/peerj-cs.623
[95] Nagelkerke, N. J. (1991). A note on a general definition of the coefficient of determination. *Biometrika*, *78*(3), 691-692. https://doi.org/10.1093/biomet/78.3.691
[96] Bergstra, J. & Bengio, Y. (2012). Random search for hyper-parameter optimization. *Journal of Machine Learning Research*, *13*, 281-305

Contact information:

Deniz AKBAY, Associate Professor
(Corresponding author)
Department of Mining and Mineral Extraction,
Çan Vocational School,
Çanakkale Onsekiz Mart University,
17020, Çanakkale, Türkiye
E-mail: denizakbay@comu.edu.tr

Gökhan EKİNCİOĞLU, Associate Professor
Department of Mining and Mineral Extraction,
Kaman Vocational School,
Kırşehir Ahi Evran University,
40100, Kırşehir, Türkiye
E-mail: gekincioglu@ahievran.edu.tr

Murat IŞIK, Assistant Professor
Department of Computer Engineering,
Faculty of Engineering and Architecture,
Kırşehir Ahi Evran University,
40100, Kırşehir, Türkiye
E-mail: muratisik@ahievran.edu.tr

Mehmet Ali YALÇINKAYA, Assistant Professor
Department of Computer Engineering,
Faculty of Engineering and Architecture,
Kırşehir Ahi Evran University,
40100, Kırşehir, Türkiye
E-mail: mehmetyalcinkaya@ahievran.edu.tr