applied sciences

MDPI

Article

GEL-LightGBM: A Gated Empirical-Law Framework for Interpretable Prediction of Rock Uniaxial Compressive Strength

Jie Peng [1], Hengyu Liu [2,*], Yun Lin [2,*] and Yanyan He [1]

[1] College of Safety Science and Engineering, Henan Polytechnic University, Jiaozuo 454003, China; 112001010008@home.hpu.edu.cn (J.P.); 112201010009@home.hpu.edu.cn (Y.H.)
[2] School of Resources and Safety Engineering, Central South University, Changsha 410083, China
* Correspondence: 19891549715@163.com (H.L.); yunlin617@csu.edu.cn (Y.L.)

Abstract

Uniaxial compressive strength (UCS) is a fundamental parameter for rock engineering design and stability assessment, but direct laboratory testing is costly, time-consuming, and often difficult for weak or fractured rocks. To improve predictive accuracy while preserving mechanical interpretability, this study proposes a Gated Empirical-Law LightGBM model (GEL-LightGBM). The framework embeds three representative rock-strength priors, including point-load strength, multi-index strength, and porosity-degradation relationships, as empirical-law experts. A sample-adaptive gating mechanism dynamically assigns its contributions for different rock states, while a controlled residual corrector captures nonlinear deviations between empirical estimates and measured UCS. Using 344 published rock-mechanics samples, porosity, Schmidt rebound hardness, P-wave velocity, and point-load strength index were used as predictors. GEL-LightGBM outperformed LightGBM, XGBoost, random forest, MLP, CNN, SVR, and BPNN, achieving a testing R2 of 0.9790 and an RMSE of 7.5623 MPa. SHAP analysis identified porosity as the dominant factor, contributing 49.0%, followed by rebound hardness (32.2%) and P-wave velocity (17.2%). The strongest interaction occurred between porosity and rebound hardness (2.31 MPa). These findings indicate that GEL-LightGBM provides accurate, stable, and physically interpretable UCS prediction for heterogeneous rock datasets.

Keywords: uniaxial compressive strength; GEL-LightGBM; empirical-law fusion; physics-guided machine learning; SHAP interpretability analysis

Check for updates

Academic Editor: Nikolaos Koukouzas

Received: 5 June 2026
Revised: 19 June 2026
Accepted: 20 June 2026
Published: 24 June 2026

Copyright: © 2026 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.

1. Introduction

Uniaxial compressive strength (UCS) is one of the most fundamental and critical mechanical parameters in rock mechanics and geotechnical engineering. It plays an important role in tunnel excavation, underground space construction, mining, slope stability analysis, foundation design, and the optimization of rock mass support schemes [1,2]. UCS not only directly reflects the load-bearing capacity of rock under axial compressive loading, but also serves as an important basis for rock mass classification, failure risk assessment, and engineering structural design [1]. As engineering construction increasingly extends to deep, complex, and strongly disturbed geological environments, rapid, accurate, and stable prediction of rock UCS has become an important research issue in rock mechanics.

At present, rock UCS is mainly obtained directly through standard laboratory uniaxial compression tests. This method provides reliable results, but it has strict requirements for specimen size, end-face preparation accuracy, loading equipment, and test process control.

Moreover, it is a destructive test with high cost, long duration, and limited applicability for field-scale implementation. For weak, fractured, weathered, or jointed rock masses, intact standard specimens are often difficult to obtain, which restricts the practical applicability of direct testing in engineering projects [1,2]. In addition, rock strength is affected by specimen shape and size effects; the height-to-width ratio, specimen scale, and distribution of internal defects can all alter the measured compressive strength [3,4]. Therefore, relying solely on standard compression tests is insufficient to meet the demand for large-scale and rapid UCS evaluation in complex engineering scenarios.

To reduce testing costs and improve engineering applicability, researchers have extensively explored UCS prediction methods based on indirect indices. Easily obtainable indicators such as point-load strength index, Schmidt rebound hardness, P-wave velocity, Leeb hardness, and porosity are commonly used to establish empirical UCS prediction relationships [3–6]. Jamshidi [7] compared the performance of different point load index test procedures in predicting sandstone UCS. Skentou et al. [8] developed a closed-form equation for predicting granite UCS based on three non-destructive testing indices and soft computing models. Kong et al. [5] systematically analyzed the correlation, validity, and reliability between index tests and UCS for building stones. Lobarinhas et al. [6] discussed the reliability of Leeb hardness as a UCS predictor for stones exposed to high temperatures and pointed out that porosity affects strength prediction across different carbonate lithologies. These studies indicate that indirect indices can reflect rock strength characteristics to some extent. However, traditional empirical equations are usually dependent on specific lithologies, limited samples, or particular testing conditions, and are prone to prediction bias when applied across different lithologies, regions, and data sources.

The variation in rock UCS is essentially controlled by multi-scale structures and multi-factor coupling effects. Setayeshirad et al. [9] conducted statistical analysis and artificial neural network modeling for carbonate rocks, showing that differences in the physical properties of carbonate rock samples can affect UCS prediction performance. Hussain et al. [10] predicted the UCS of dolerite based on petrographic features and a hybrid GMDH-LM model, further demonstrating the importance of microstructural information for strength prediction. In addition, freeze–thaw damage [11], hydro-chemical corrosion [12], weathering state [13], and mineral micromechanical properties [14] can also change the strength response of rocks. These findings suggest that rock strength response is not linearly determined by a single indicator but is jointly controlled by factors such as pore defects, mineral structure, surface hardness, internal integrity, and environmental actions.

In recent years, machine learning methods have provided a new technical route for UCS prediction. Compared with traditional empirical equations, machine learning models can automatically learn complex nonlinear mapping relationships from multi-source data and have shown high accuracy in rock strength prediction. Piao et al. [15] proposed a hybrid Bayesian group-based machine learning method for predicting rock UCS, improving the robustness of model prediction. Liu et al. [16] developed stacking models based on lithological features and demonstrated that ensemble learning can effectively enhance UCS prediction accuracy. Khatti and Grover [17] used a relevance vector machine with dual kernels combined with metaheuristic algorithms to predict rock uniaxial strength. Zhao et al. [18] proposed a fully Bayesian Gaussian process regression model and improved prediction reliability under limited sample conditions through model class selection. Meanwhile, Davoodi et al. [19] and Zhao et al. [20] predicted UCS using drilling variables and measurement-while-drilling data, respectively, promoting the transition of strength prediction from laboratory post-analysis toward real-time field evaluation.

Despite the significant progress achieved by machine learning methods, existing models still have several limitations. First, purely data-driven models are highly dependent

on the size, representativeness, and distribution characteristics of training samples, and they are prone to insufficient generalization under small-sample, multi-lithology, and multi-source heterogeneous data conditions [15,18,21]. Second, although conventional machine learning models can improve prediction accuracy, they often lack explicit rock mechanics constraints. As a result, their prediction processes are difficult to interpret, making it challenging for engineers to judge whether the model conforms to basic strength evolution laws. Third, traditional empirical equations have clear physical meanings but limited expressive capacity, whereas machine learning models have strong expressive ability but may weaken mechanical priors. Establishing an effective connection between empirical laws and data-driven learning has therefore become a key issue for improving the accuracy, stability, and interpretability of UCS prediction. Studies combining dimensional analysis with machine learning have shown that incorporating physical knowledge into intelligent models can improve generalization performance [22]. However, further research is still needed on how to adaptively fuse classical empirical laws in UCS prediction, how to correct the bias of empirical equations, and how to interpret the coupled control mechanisms of multiple indices.

To address these issues, this study proposes a Gated Empirical-Law LightGBM model (GEL-LightGBM) for UCS prediction under multi-source rock sample conditions. The model uses porosity, Schmidt rebound hardness, P-wave velocity, and point-load strength index as input variables, with UCS as the prediction target. The relationships of point-load strength, multi-index comprehensive strength, and porosity degradation are organized as empirical-law experts within the model. On this basis, the model dynamically assigns the contributions of different empirical laws under different rock states through a sample-adaptive gating mechanism and uses a controlled residual corrector to characterize the nonlinear deviation between empirical priors and measured strength. In this way, the empirical knowledge of rock mechanics is synergistically integrated with the nonlinear learning capability of LightGBM.

The main contributions of this study are as follows:

(1) An empirical-law expert layer integrating point-load strength, multi-index comprehensive strength, and porosity-degradation relationships is constructed. This enables the model to explicitly use existing empirical laws of rock strength and avoids the complete dependence of purely data-driven models on sample distributions.

(2) A sample-adaptive gated fusion mechanism is proposed. Based on the porosity, Schmidt rebound hardness, P-wave velocity, and point-load strength index of different rock samples, the mechanism dynamically adjusts the weights of each empirical-law expert, thereby improving the adaptability of the model to multi-source heterogeneous rock data.

(3) A controlled residual correction strategy is introduced. While preserving the prior constraints of empirical laws, LightGBM is used to characterize the nonlinear deviation between empirical predictions and measured UCS, thereby improving prediction accuracy and generalization stability.

(4) SHAP-based interpretability analysis is combined to reveal the contribution direction, importance ranking, and interaction effects of porosity, Schmidt rebound hardness, P-wave velocity, and point-load strength index on UCS prediction. This provides a basis for understanding the coupled control mechanism of pore defects, surface hardness, and internal integrity on rock compressive strength.

The remainder of this paper is organized as follows. Section 2 introduces the database sources, statistical characteristics of variables, correlation analysis, and data preprocessing methods. Section 3 describes the basic LightGBM model and the empirical-law layer, gated fusion layer, residual correction layer, and training strategy of GEL-LightGBM. Section 4 validates the effectiveness of the proposed method through model comparison,

ablation experiments, and interpretability analysis. Section 5 discusses the generalization advantages, physical interpretability, and applicability boundaries of the model. Section 6 summarizes the main conclusions and outlines future research directions.

2. Database

2.1. Data Sources and Description

The database used for uniaxial compressive strength (UCS) prediction in this study was collected from publicly published literature in rock mechanics and geotechnical engineering. The data were compiled from the studies of Mahmoodzadeh et al. [23], Kilic and Teymen [24], Dehghan et al. [25], Armaghani et al. [26], and Aladejare [27], covering various lithologies, including granite, limestone, sandstone, tuff, and travertine, as well as experimental samples from different regions. Therefore, the database is not limited to rocks from a single engineering site or a specific geographic area, but represents a multi-source collection of commonly reported rock types with different physical and mechanical characteristics. The database uses porosity (n), Schmidt rebound hardness (SHR), P-wave velocity (V_p), and point-load strength index $I_s(50)$ as input variables, with uniaxial compressive strength (UCS) as the prediction target. After unifying variable names, units, and data formats, a total of 344 complete samples were retained for model training and analysis. This database integrates direct UCS test results with commonly used indirect index test results, providing a multi-source data basis for developing UCS prediction models based on physical and mechanical indices.

Figure 1 presents the violin plots, box plots, and observed scatter distributions of all variables. Overall, the value ranges and distribution patterns differ markedly among the variables, reflecting the heterogeneity of the multi-source rock samples in terms of pore structure, surface hardness, internal integrity, and strength level. Specifically, porosity (n) is mainly concentrated in the lower range, while a small number of high-porosity samples are also present. This results in a distinctly right-skewed distribution with a sharp peak and heavy tail, consistent with the high skewness (Skew = 2.846) and kurtosis (Kurt = 11.785) reported in Table 1. The SHR values range from 17.00 to 67.07, with a relatively wide box distribution, indicating considerable variation in surface hardness among the samples. Its mean value is lower than the median, suggesting that the distribution is slightly skewed toward the lower-hardness range. V_p is mainly concentrated around 5000–6000 m/s, but its overall range spans from 1470 to 7943 m/s, indicating significant differences in internal compactness and fracture development among the samples. $I_s(50)$ is mostly concentrated in the low-to-medium strength range, with a small number of high-value samples, showing a certain degree of right-skewness. UCS has the widest value range, from 6.20 to 239.00 MPa, indicating that the database contains both low-strength and high-strength rock samples, which is beneficial for evaluating the predictive capability of the model across different strength levels.

The above distributional characteristics indicate that the database is not limited to a single lithology or a single strength range, but instead exhibits clear multi-source, dispersed, and imbalanced distribution characteristics. In particular, the skewed distributions and high-value samples observed in porosity, $I_s(50)$, and UCS increase the fitting difficulty for both traditional empirical equations and purely data-driven models. Therefore, in the subsequent modeling process, it is necessary to adopt a prediction method capable of simultaneously handling nonlinear relationships, differences in variable scales, and sample heterogeneity.

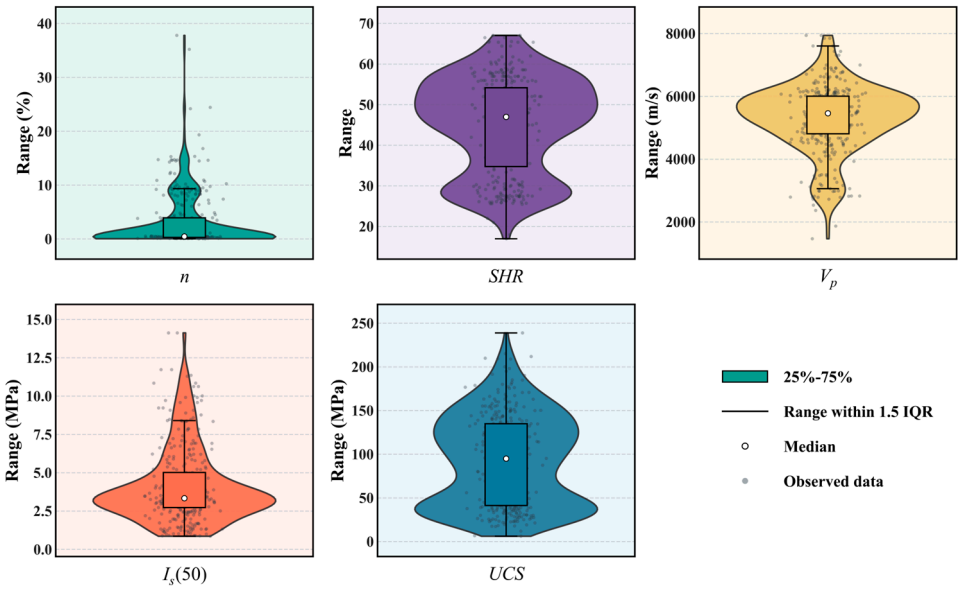


Figure 1. Distribution of the dataset.

Table 1. Statistical characteristics of the variables.

Variable	Mean	Median	Std	Kurt	Skew	Min	Max
Is(50)	3.522	3.340	2.466	1.819	1.361	0.860	14.130
SHR	43.271	47.000	11.733	−0.990	−0.303	17.000	67.070
Vp	5181.491	5463.000	1085.015	0.833	−0.688	1470.000	7943.000
n	0.908	0.470	5.091	11.785	2.846	0.060	37.810
UCS	76.032	95.070	52.261	−0.983	0.259	6.200	239.000

2.2. Correlation Analysis

To reveal the degree of association between each physical-mechanical parameter and uniaxial compressive strength (UCS), the Pearson correlation coefficient was used in this study to analyze the correlations among porosity (n), Schmidt rebound hardness (SHR), P-wave velocity (V_p), point-load strength index $I_s(50)$, and UCS. The Pearson correlation coefficient [28] is calculated as follows:

$$r_{xy} = \frac{\sum_{i=1}^{n}(x_i - \overline{x})(y_i - \overline{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \overline{x})^2}\sqrt{\sum_{i=1}^{n}(y_i - \overline{y})^2}} \quad (1)$$

where r_{xy} is the correlation coefficient between variables x and y, with a value range of $-1, 1$; $r > 0$ indicates a positive correlation, whereas $r < 0$ indicates a negative correlation. The closer $|r|$ is to 1, the stronger the correlation.

In Figure 2, *, **, and *** indicate statistically significant Pearson correlations at the $p < 0.05$, $p < 0.01$, and $p < 0.001$ levels, respectively. As shown in Figure 2, SHR exhibits the strongest correlation with UCS, with a correlation coefficient of $r = 0.75$, indicating that rebound hardness is an important indicator for characterizing rock compressive strength. This is mainly because SHR reflects the ability of the rock surface to resist impact and local deformation. In general, denser rocks with harder mineral particles and higher cementation degrees tend to have higher rebound values and, correspondingly, higher UCS. V_p also shows a strong positive correlation with UCS, with a correlation coefficient of $r = 0.61$. From a mechanistic perspective, V_p is controlled by internal pores, fractures, and

structural integrity. Rocks with more intact structures and fewer defects allow faster elastic wave propagation and usually have a more stable load-bearing skeleton, leading to higher compressive strength. Porosity (n) shows a clear negative correlation with UCS, with a correlation coefficient of $r = -0.58$, indicating that pore development weakens the effective load-bearing area and promotes stress concentration and crack propagation during loading, thereby reducing rock strength. In comparison, $I_s(50)$ has a moderate correlation with UCS, with a correlation coefficient of $r = 0.41$. This suggests that the point-load strength index can reflect the failure resistance of rock to some extent, but its standalone characterization capability is weaker than that of SHR and V_p because it is affected by specimen shape, loading mode, and lithological differences.



Figure 2. Correlation heatmap.

Overall, the correlation analysis shows that SHR, V_p, n, and $I_s(50)$ are all clearly correlated with UCS. These variables can reflect rock compressive strength characteristics from different perspectives, including surface hardness, internal integrity, pore structure, and local failure strength. Among them, SHR and V_p have stronger characterization capabilities for UCS, while porosity shows a distinct weakening effect. Meanwhile, the scatter distributions indicate that the relationships among some variables are not strictly linear, suggesting that UCS is jointly controlled by multiple factors and involves certain nonlinear coupling characteristics. Therefore, relying solely on a single empirical relationship is insufficient to fully describe the variation pattern of rock strength, and it is necessary to further use machine learning models to comprehensively explore the complex mapping relationships between multiple parameters and UCS.

2.3. Data Preprocessing

To ensure the completeness of the database and the stability of model training, missing values in the original data were first checked. Since only a small amount of missing data existed in the database, the K-nearest neighbor imputation (KNN) method [29] was used for data filling. This method is based on the similarity between samples. It identifies the k

samples closest to the sample with missing values and estimates the missing entries using the corresponding variable values of the neighboring samples. Compared with mean or median imputation, KNN imputation can comprehensively consider the relationships among multiple variables, thereby recovering missing information more reasonably. Because the number of missing values was small, the overall distribution characteristics of each variable did not change significantly after imputation.

To eliminate the influence of differences in dimensions and numerical scales among different input variables on model training, the input features were normalized in this study [30]. The normalization formula is expressed as:

$$x' = \frac{x - x_{\min}}{x_{\max} - x_{\min}} \tag{2}$$

where x is the original feature value, x' is the normalized feature value, and $x_{\min}$ and $x_{\max}$ are the minimum and maximum values of the corresponding feature, respectively. After normalization, all input variables were mapped into a unified numerical range, which helps improve model training efficiency and prevents variables with larger numerical scales from occupying excessive weights during model learning. Finally, the preprocessed data were used for the construction and evaluation of subsequent UCS prediction models. For model evaluation, the 344 samples were randomly divided into training and testing sets at a ratio of 80:20. The same data partition was used for GEL-LightGBM, all benchmark models, and the ablation variants to ensure a consistent comparison. The model implementations followed the default random-state setting of the Python modeling environment, and no additional manual re-partitioning was introduced during model comparison.

3. Method

To achieve high-accuracy prediction of rock uniaxial compressive strength and enhance the model's ability to utilize classical empirical laws of rock strength, this study proposes a Gated Empirical-Law LightGBM model, namely GEL-LightGBM. The model uses porosity (n), Schmidt rebound hardness (SHR), P-wave velocity (V_p), and point-load strength index $I_s(50)$ as input variables, with uniaxial compressive strength (UCS) as the prediction target. GEL-LightGBM does not simply use the calculated results of empirical equations as additional input features. Instead, it organizes classical empirical UCS relationships as empirical-law experts within the model and dynamically learns the reliability of different empirical laws under different rock states through a sample-adaptive gating mechanism. Subsequently, the model further employs a controlled residual corrector to characterize the nonlinear deviation between the empirical-law priors and the measured strength.

3.1. LightGBM Model

LightGBM is an efficient ensemble learning algorithm based on gradient boosting decision trees, proposed by Ke et al. [31]. Given a training sample set:

$$D = \{(\mathbf{x}_i, y_i)\}_{i=1}^{N} \tag{3}$$

where $\mathbf{x}_i$ denotes the input variables of the i-th sample and y_i denotes the corresponding measured UCS, the prediction function of LightGBM can be expressed as:

$$\hat{y}_i = \sum_{m=1}^{M} f_m(\mathbf{x}_i) \tag{4}$$

where M is the number of regression trees and $f_m(\cdot)$ represents the m-th regression tree.

The model is trained by minimizing a regularized objective function:

$$\mathcal{L}_{\mathrm{LGBM}} = \sum_{i=1}^{N} l(y_i, \hat{y}_i) + \sum_{m=1}^{M} \Omega(f_m) \tag{5}$$

where $l(\cdot)$ is the prediction loss function and $\Omega(f_m)$ is the penalty term for tree structure complexity. LightGBM can effectively handle nonlinear relationships and variable interactions, making it suitable for tabular UCS prediction tasks with limited rock mechanics parameters and complex variable relationships. However, conventional LightGBM relies entirely on data distributions for fitting and does not explicitly incorporate empirical rock strength relationships such as point-load strength, wave velocity response, rebound hardness, and porosity degradation. Therefore, this study further develops GEL-LightGBM by introducing empirical-law embedding, sample-level gated selection, and residual correction mechanisms into the LightGBM framework.

3.2. GEL-LightGBM Model Framework

The overall prediction form of GEL-LightGBM is given by:

$$\widehat{UCS}_i = UCS_{\mathrm{phy},i} \cdot \exp(\eta R_{\mathrm{LGBM}}(\mathbf{x}_i)) \tag{6}$$

where $\mathbf{x}_i = [n_i, SHR_i, V_{p,i}, I_{s,i}(50)]$, $UCS_{\mathrm{phy},i}$ is the empirical strength prior obtained from gated empirical-law fusion, $R_{\mathrm{LGBM}}(\mathbf{x}_i)$ is the logarithmic residual output by the residual corrector, and η is the residual shrinkage coefficient satisfying $0 < \eta \leq 1$. The coefficient η is used to control the correction magnitude of the data-driven residual on the empirical-law prior, thereby preventing an excessively strong residual model from reducing the empirical-law layer to a merely formal component.

3.2.1. Empirical-Law Layer

The empirical-law layer is used to embed classical empirical relationships of rock strength into GEL-LightGBM. To avoid model redundancy caused by using an excessive number of equations, three representative types of empirical laws are selected in this study, corresponding to the point-load strength prior, the multi-index comprehensive strength prior, and the porosity-degradation prior, respectively. Unless otherwise specified, the units of UCS and $I_s(50)$ are MPa, and the unit of V_p is km/s. In the calculation of the empirical equations, V_p is converted from m/s to km/s before substitution. Porosity (n) is input in percentage form.

(1) Point-load strength prior:

$$F_1 = 24 I_s(50) \tag{7}$$

The point-load strength index $I_s(50)$ has a direct strength correlation with UCS. Broch and Franklin [32] and the ISRM suggested method [33] indicated that the point-load test can be used to indirectly estimate rock uniaxial compressive strength, and the commonly used conversion coefficient generally falls within the range of 20–25. Therefore, F_1 is used as the basic strength, dominated prior by point-load strength.

(2) Multi-index comprehensive empirical law:

$$F_2 = 4.14 I_s(50) + 29.8 V_p + 0.54 SHR - 116 \tag{8}$$

This equation is derived from the multiple regression relationship proposed by Çobanoğlu and Çelik [34]. This empirical law simultaneously includes $I_s(50)$, V_p, and SHR, which characterize local strength, rock compactness, and surface hardness, respec-

tively. Therefore, it has a high degree of consistency with the input variables used in this study.

(3) Porosity-degradation prior:

$$F_3 = 147.16 \exp(-0.0835n) \tag{9}$$

This equation is derived from the porosity-UCS relationship established by Kılıç and Teymen [24] based on 19 rock types. Its exponential decay form reflects the basic tendency that increasing porosity weakens the continuity of the rock skeleton, increases potential fracture space, and reduces load-bearing capacity. Therefore, it is suitable as the porosity-degradation expert in the proposed model.

Because the above empirical equations originate from different lithologies, testing conditions, and data ranges, direct use may introduce scale bias. In addition, because F_2 contains an intercept term, non-positive values may occur for low-strength samples. To ensure the positive definiteness of the empirical-law outputs with respect to strength, a smooth positive projection is first applied:

$$F_j^+ = F_{\min} + \alpha_p \log[1 + \exp((F_j - F_{\min})/\alpha_p)] \tag{10}$$

where $F_{\min}$ is the minimum physical strength threshold, which can be set to 1 MPa, and α_p is the smoothing parameter for the projection.

Subsequently, each empirical-law expert is calibrated on a logarithmic scale:

$$E_j = \exp(a_j + b_j \log F_j^+) \tag{11}$$

where a_j and b_j are calibration parameters. The optimization objective is defined as:

$$\mathcal{L}_{\text{cal}} = \sum_{j=1}^{3} \sum_{i=1}^{N} \rho\left(\log y_i - a_j - b_j \log F_{ij}^+\right) \tag{12}$$

where $\rho(\cdot)$ can be the squared loss or Huber loss. This calibration process is used to correct systematic offsets among different empirical equations caused by differences in applicable lithologies and testing conditions. After positive projection and calibration, E_1, E_2, and E_3 are used as the effective outputs of the three empirical-law experts, respectively.

3.2.2. Gated Fusion Layer

Different empirical equations do not have the same reliability for different rock samples. Fixed-weight fusion assumes that all samples share the same set of contribution coefficients for the empirical equations, which makes it difficult to reflect sample heterogeneity caused by differences in pore structure, wave velocity response, rebound hardness, and point-load strength. Therefore, this study designs a sample-adaptive gated fusion layer, in which the empirical-law weights are dynamically determined by the state of the current sample.

Considering that a conventional LightGBM regressor is a single-output model, three parallel gated LightGBM submodels are adopted to learn the reliability scores of the three empirical-law experts:

$$s_{ij} = G_j(\mathbf{x}_i) \tag{13}$$

where $G_j(\cdot)$ denotes the j-th gating submodel, and s_{ij} is the unnormalized score of the j-th empirical-law expert for the i-th sample. The sample-level gating weights are then obtained using the Softmax function:

$$w_{ij} = \frac{\exp(s_{ij}/T)}{\sum_{k=1}^{3} \exp(s_{ik}/T)} \tag{14}$$

where T is the temperature parameter. The gating weights satisfy:

$$\sum_{j=1}^{3} w_{ij} = 1, \quad w_{ij} \geq 0 \tag{15}$$

To train the gating models, soft reliability labels are first constructed based on the empirical-law errors. The local logarithmic error of the j-th empirical-law expert for the i-th sample is defined as:

$$e_{ij} = \left|\log y_i - \log E_{ij}\right| \tag{16}$$

A smaller error indicates that the empirical-law expert is more reliable for the current sample. Therefore, the soft reliability label is defined as:

$$q_{ij} = \frac{\exp(-e_{ij}/\tau)}{\sum_{k=1}^{3} \exp(-e_{ik}/\tau)} \tag{17}$$

where τ is the smoothing coefficient. The training objective of the three gating submodels is:

$$\mathcal{L}_{\text{gate}} = \sum_{i=1}^{N} \sum_{j=1}^{3} \left[G_j(\mathbf{x}_i) - \log(q_{ij} + \epsilon_q)\right]^2 \tag{18}$$

where ϵ_q is a small constant used to prevent numerical instability.

It should be emphasized that w_{ij} is a sample-dependent function rather than a fixed constant shared by the entire dataset. For each rock sample, GEL-LightGBM reallocates the contributions of the three empirical-law experts according to $\mathbf{x}_i$. The gated empirical-law fusion is performed in logarithmic space:

$$UCS_{\text{phy},i} = \exp\left(\sum_{j=1}^{3} w_{ij} \log E_{ij}\right) \tag{19}$$

This expression is equivalent to a sample-adaptive weighted geometric fusion, which is suitable for positive-valued strength parameters and can reduce the dominance of abnormal outputs from any single empirical equation on the fusion result.

3.2.3. Residual Correction Layer and Overall Objective Function

The UCS_{phy} obtained through gated empirical-law fusion can provide a strength prior with rock mechanics significance. However, empirical equations still cannot fully represent complex factors such as rock heterogeneity, differences in mineral composition, microcrack development, and disturbances caused by testing conditions. Therefore, a residual correction layer is further introduced in this study to model the proportional deviation between the empirical-law fusion prior and the measured UCS.

First, the logarithmic residual is defined as:

$$r_i = \log\left(\frac{y_i}{UCS_{\text{phy},i}}\right) \tag{20}$$

Then, a residual LightGBM model is trained as:

$$R_{\mathrm{LGBM}}(\mathbf{x}_i) \approx r_i \tag{21}$$

The optimization objective of the residual corrector is:

$$\mathcal{L}_{\mathrm{res}} = \sum_{i=1}^{N} [r_i - R_{\mathrm{LGBM}}(\mathbf{x}_i)]^2 \tag{22}$$

The final predicted value is defined as:

$$\widehat{UCS}_i = UCS_{\mathrm{phy},i} \cdot \exp(\eta R_{\mathrm{LGBM}}(\mathbf{x}_i)) \tag{23}$$

The corresponding final prediction loss can be written as:

$$\mathcal{L}_{\mathrm{pred}} = \sum_{i=1}^{N} \left[\log y_i - \log \widehat{UCS}_i\right]^2 \tag{24}$$

Therefore, the overall objective function of GEL-LightGBM can be expressed as:

$$\mathcal{L}_{\mathrm{GEL}} = \lambda_{\mathrm{cal}}\mathcal{L}_{\mathrm{cal}} + \lambda_{\mathrm{gate}}\mathcal{L}_{\mathrm{gate}} + \lambda_{\mathrm{res}}\mathcal{L}_{\mathrm{res}} + \lambda_{\mathrm{pred}}\mathcal{L}_{\mathrm{pred}} \tag{25}$$

where λ_{cal}, λ_{gate}, λ_{res}, and λ_{pred} are the weights of the different loss terms. It should be noted that, because LightGBM is a tree-based model, this study does not adopt neural-network-style end-to-end joint optimization through backpropagation. Instead, the above overall objective is implemented in a staged manner: the empirical-law experts are first calibrated, the gating submodels are then trained, and finally the residual corrector is trained and the final prediction form is determined. The model architecture is shown in Figure 3.

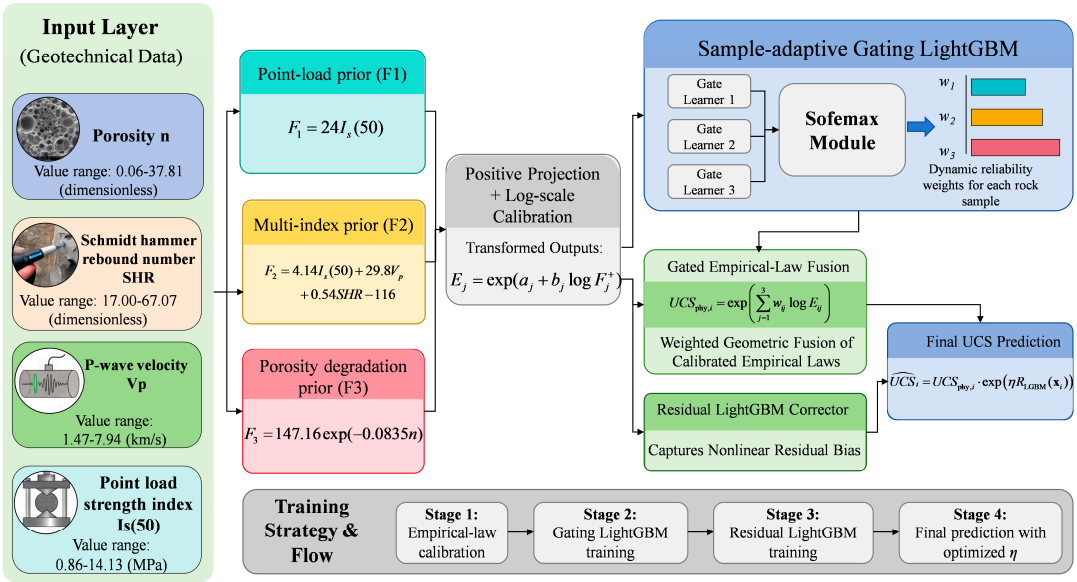


Figure 3. Architecture of the GEL-LightGBM model.

3.3. Training Strategy of GEL-LightGBM

GEL-LightGBM is trained using a staged cross-fitting strategy to maintain the functional separation among empirical-law calibration, gated fusion, and residual correction. First, the three empirical-law outputs F_1, F_2, and F_3 are calculated on the training data,

and E_1, E_2, and E_3 are obtained through positive projection and logarithmic calibration. Then, soft reliability labels q_{ij} are constructed based on the local logarithmic errors of each empirical-law expert on the samples, and three parallel gated LightGBM submodels, G_1, G_2, and G_3, are trained to generate sample-level empirical-law weights.

To improve the robustness of the residual correction target, this study adopts a K-fold cross-fitting strategy to generate the empirical-law fusion prior for the training samples. It should be noted that this cross-fitting procedure was performed only within the training set. The testing set was not involved in empirical-law calibration, gated LightGBM training, residual target construction, or residual shrinkage coefficient determination, and was kept independent for final performance evaluation. Specifically, in each iteration, $K-1$ subsets are used to train the empirical-law calibrators and gating submodels, and the remaining subset is used to predict $UCS_{\mathrm{phy},i}^{\mathrm{OOF}}$. After repeating this process, each training sample obtains a cross-fitted empirical prior prediction. The logarithmic residual is then constructed as:

$$r_i^{\mathrm{OOF}} = \log\left(\frac{y_i}{UCS_{\mathrm{phy},i}^{\mathrm{OOF}}}\right) \tag{26}$$

The residual LightGBM is trained using $\mathbf{x}_i$ as the input and r_i^{OOF} as the target.

The residual shrinkage coefficient η is determined based on the cross-fitted prediction results:

$$\eta^* = \mathrm{argmin}_{0<\eta\le 1}\sum_{i=1}^{N}\left[\log y_i - \log UCS_{\mathrm{phy},i}^{\mathrm{OOF}} - \eta R_{\mathrm{LGBM}}^{\mathrm{OOF}}(\mathbf{x}_i)\right]^2 \tag{27}$$

This one-dimensional optimization can be solved by grid search or truncated least squares. During final prediction, the empirical-law calibrators, gating submodels, and residual corrector are refitted using the entire training dataset. For a sample to be predicted, the model sequentially calculates the empirical-law outputs, sample-level gating weights, empirical fusion prior, and residual correction term, and then obtains the final prediction:

$$\widehat{UCS} = UCS_{\mathrm{phy}} \cdot \exp(\eta^* R_{\mathrm{LGBM}}(\mathbf{x})) \tag{28}$$

For clarity and reproducibility, the complete training and prediction procedure of GEL-LightGBM is summarized in Algorithm 1.

Algorithm 1: Training and prediction procedure of GEL-LightGBM

Input:
Training dataset $D = \{(\mathbf{x}_i, y_i)\}_{i=1}^{N}$, where $\mathbf{x}_i = [n_i, SHR_i, V_{p,i}, I_{s,i}(50)]$; test samples $\mathbf{x}_{test}$; number of cross-fitting folds K.

Output:
Predicted UCS values $\widehat{UCS}$.

1. Calculate the empirical-law outputs F_1, F_2, and F_3 using Equations (7)–(9).
2. Apply positive projection to the empirical-law outputs and obtain F_j^+.
3. Calibrate each empirical-law expert on the logarithmic scale to obtain E_1, E_2, and E_3.
4. Compute the local logarithmic error $e_{ij} = \left|\log y_i - \log E_{ij}\right|$ for each empirical-law expert.
5. Construct the soft reliability label q_{ij} according to Equation (17).
6. Train three gated LightGBM submodels G_1, G_2, and G_3 to estimate the reliability scores of the empirical-law experts.
7. Convert the reliability scores into sample-level gating weights w_{ij} using the Softmax function.

Algorithm 1: *Cont.*

8. Fuse the calibrated empirical-law experts in logarithmic space to obtain the empirical strength prior $UCS_{phy,i}$.
9. Use K-fold cross-fitting to generate out-of-fold empirical priors $UCS_{phy,i}^{OOF}$ for the training samples.
10. Construct the logarithmic residual $r_i^{OOF} = \log(y_i / UCS_{phy,i}^{OOF})$.
11. Train the residual LightGBM corrector $R_{LGBM}(\mathbf{x}_i)$ using $\mathbf{x}_i$ as the input and r_i^{OOF} as the target.
12. Determine the residual shrinkage coefficient η^* using Equation (27).
13. Refit the empirical-law calibrators, gating submodels, and residual corrector using the full training dataset.
14. For each test sample, calculate UCS_{phy}, $R_{LGBM}(\mathbf{x})$, and the final prediction $\widehat{UCS} = UCS_{phy} \cdot \exp(\eta^* R_{LGBM}(\mathbf{x}))$.

3.4. Implementation Environment

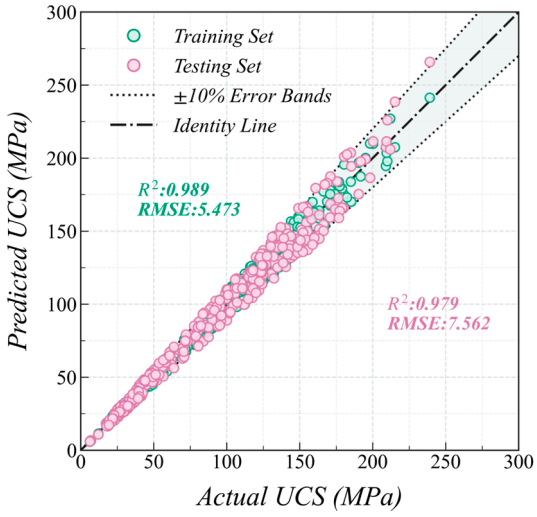
All model construction, training, testing, ablation experiments, and SHAP-based interpretability analyses were implemented using Python 3.12 in Jupyter Notebook 7.4.5 under the Windows 11 operating system. The computational experiments were carried out on a workstation equipped with an NVIDIA GeForce RTX 5080 GPU. The proposed GEL-LightGBM model, the comparison models, and the ablation variants were implemented and evaluated in the same software environment, using the same data partition, input variables, and evaluation metrics to ensure consistency across all experiments.

4. Results

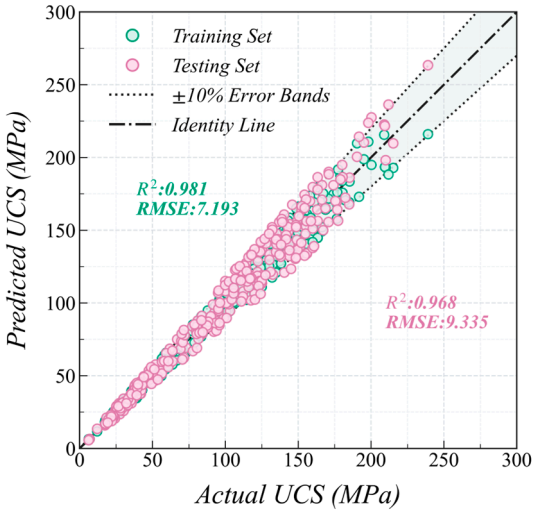
4.1. Prediction Results

To systematically evaluate the performance of the proposed GEL-LightGBM model in predicting rock uniaxial compressive strength (UCS), it was compared with seven commonly used machine learning and deep learning models, including LightGBM, XGBoost, random forest (RF), multilayer perceptron (MLP), convolutional neural network (CNN), support vector regression (SVR), and back-propagation neural network (BPNN). Each model was used to make predictions on both the training and testing sets. The coefficient of determination (R^2), root mean square error (RMSE), mean absolute percentage error (MAPE), and mean absolute error (MAE) were adopted as evaluation metrics. The scatter plots of prediction results and the histograms of residual distributions are shown in Figures 4 and 5, respectively.

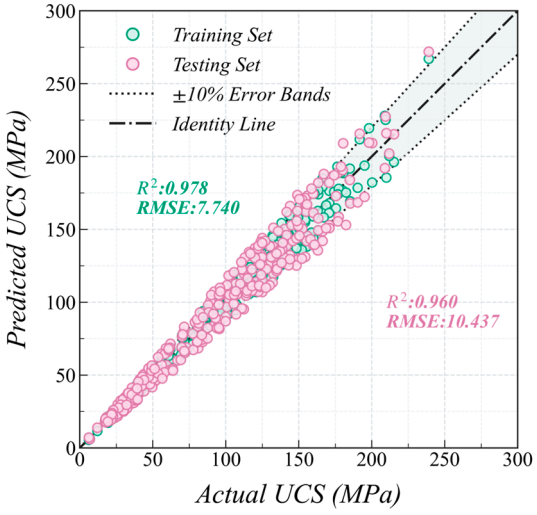
As shown in the prediction scatter plots in Figure 4, all eight models can capture the positive correlation between measured UCS and predicted UCS to some extent, but clear differences exist in prediction accuracy and generalization ability. Overall, the scatter points of GEL-LightGBM are most densely distributed near the ideal prediction line, and most samples in both the training and testing sets fall within the $\pm 10\%$ error band, indicating the best fitting accuracy and testing stability. The training R^2 and RMSE of GEL-LightGBM are 0.989 and 5.473 MPa, respectively, while the testing R^2 and RMSE are 0.979 and 7.562 MPa, respectively, outperforming the other seven models. Compared with conventional LightGBM, GEL-LightGBM reduces the testing RMSE from 9.335 MPa to 7.562 MPa, indicating that the empirical-law priors, sample-adaptive gated fusion, and residual correction mechanism can effectively improve the model's ability to represent the nonlinear variation in UCS across multiple lithologies.



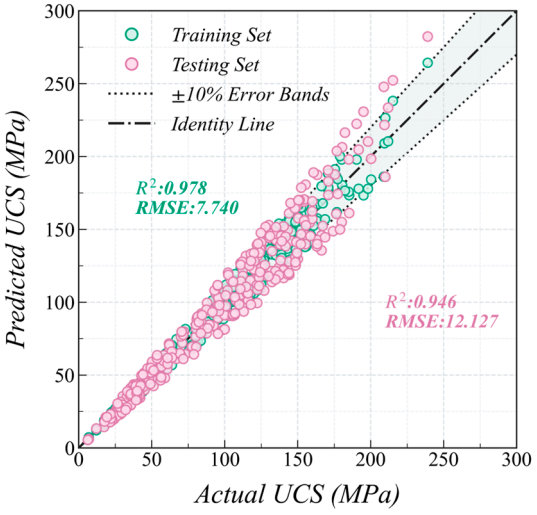
(**a**) GEL-LightGBM



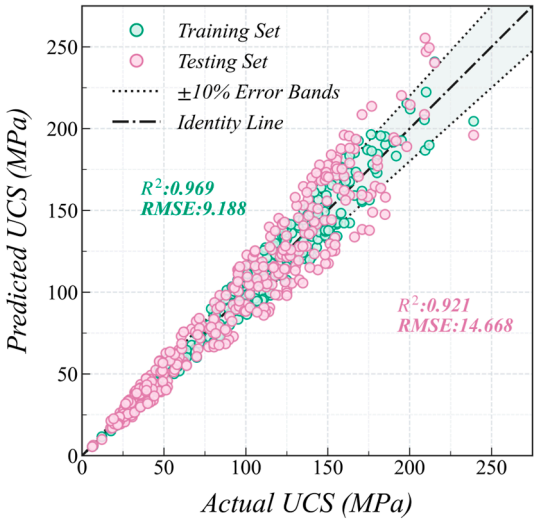
(**b**) LightGBM



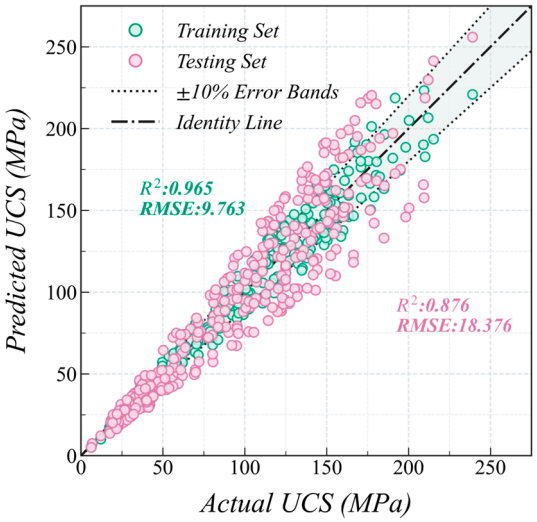
(**c**) XGBoost



(**d**) RF

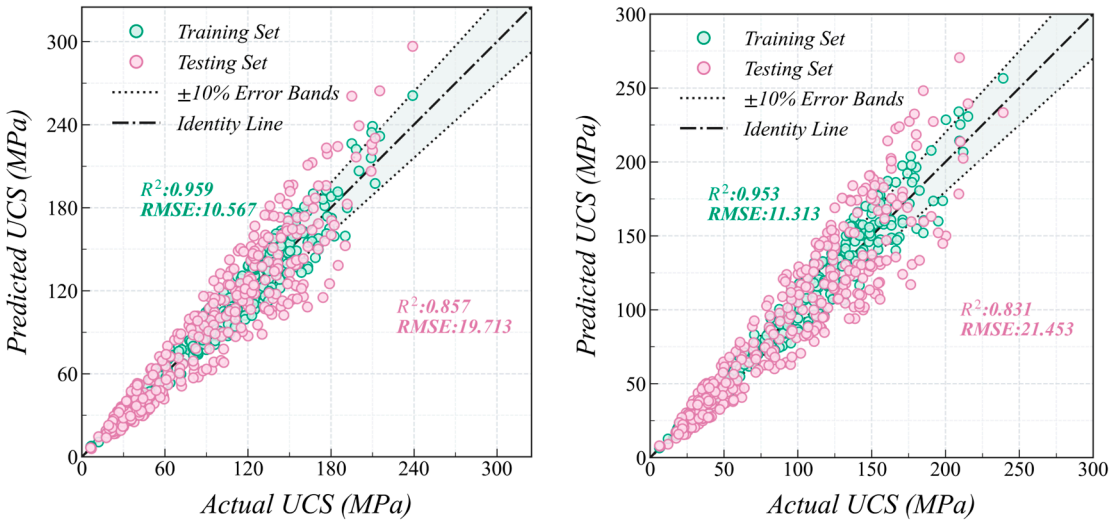


(**e**) MLP



(**f**) CNN

Figure 4. *Cont.*



(**g**) SVR (**h**) BPNN

Figure 4. Scatter plots of prediction results.

The scatter points of LightGBM, XGBoost, and RF are also mainly distributed along the ideal prediction line. However, from LightGBM to RF, the testing R^2 decreases from 0.968 and 0.960 to 0.946, while the RMSE increases from 9.335 and 10.437 to 12.127 MPa. Some medium- and high-strength samples gradually deviate from the error band, suggesting that although purely data-driven ensemble models have strong nonlinear fitting ability, they still exhibit certain fluctuations when dealing with complex lithological differences and high-strength sample prediction. In comparison, the scatter distributions of MLP, CNN, SVR, and BPNN are more dispersed. Their testing R^2 values decrease to 0.921, 0.876, 0.857, and 0.831, respectively, and the corresponding RMSE values increase to 14.668, 18.376, 19.713, and 21.453 MPa. The deviations from the ideal prediction line are particularly evident in the higher UCS range, indicating that these models have relatively insufficient generalization ability under the current limited-sample and multi-source heterogeneous rock database. Overall, GEL-LightGBM not only maintains consistent performance between the training and testing sets but also shows smaller prediction bias over the full strength range, demonstrating that the proposed empirical-law-guided gated learning framework can effectively improve the accuracy, stability, and engineering applicability of UCS prediction.

The histograms of residual distributions in Figure 5 further verify the differences in the dispersion of prediction errors and systematic bias among the models. Overall, the residual distribution of GEL-LightGBM is closest to the ideal state. The residuals of both the training and testing sets are highly concentrated around zero error and have the narrowest distribution width. The residual standard deviations of the training and testing sets are only 5.47 MPa and 7.56 MPa, respectively, and the corresponding residual means are 0.13 MPa and −0.026 MPa. This indicates that the model not only has small prediction errors, but also shows almost no obvious systematic overestimation or underestimation. In comparison, the residual distributions of LightGBM, XGBoost, and RF are still generally centered around zero, but their testing residual standard deviations gradually increase to 9.25 MPa, 10.4 MPa, and 12.1 MPa, respectively. This suggests that purely data-driven ensemble models exhibit more pronounced error fluctuations on testing samples. The residual distributions of MLP, CNN, SVR, and BPNN are noticeably broader, with testing standard deviations increasing to 14.6 MPa, 18.4 MPa, 19.7 MPa, and 21.5 MPa, respectively, and with longer distribution tails. This indicates that these models produce relatively

large prediction deviations for some samples and have insufficient generalization stability. In particular, BPNN has the widest residual distribution range and the largest standard deviation on the testing set, suggesting that it struggles to fully capture the complex nonlinear relationships between UCS and input indices for multi-lithology rocks. Overall, GEL-LightGBM performs best in terms of both a residual mean close to zero and the smallest residual standard deviation. This demonstrates that empirical-law guidance, gated fusion, and residual correction can effectively reduce the dispersion of prediction errors and improve the stability and reliability of UCS prediction results.

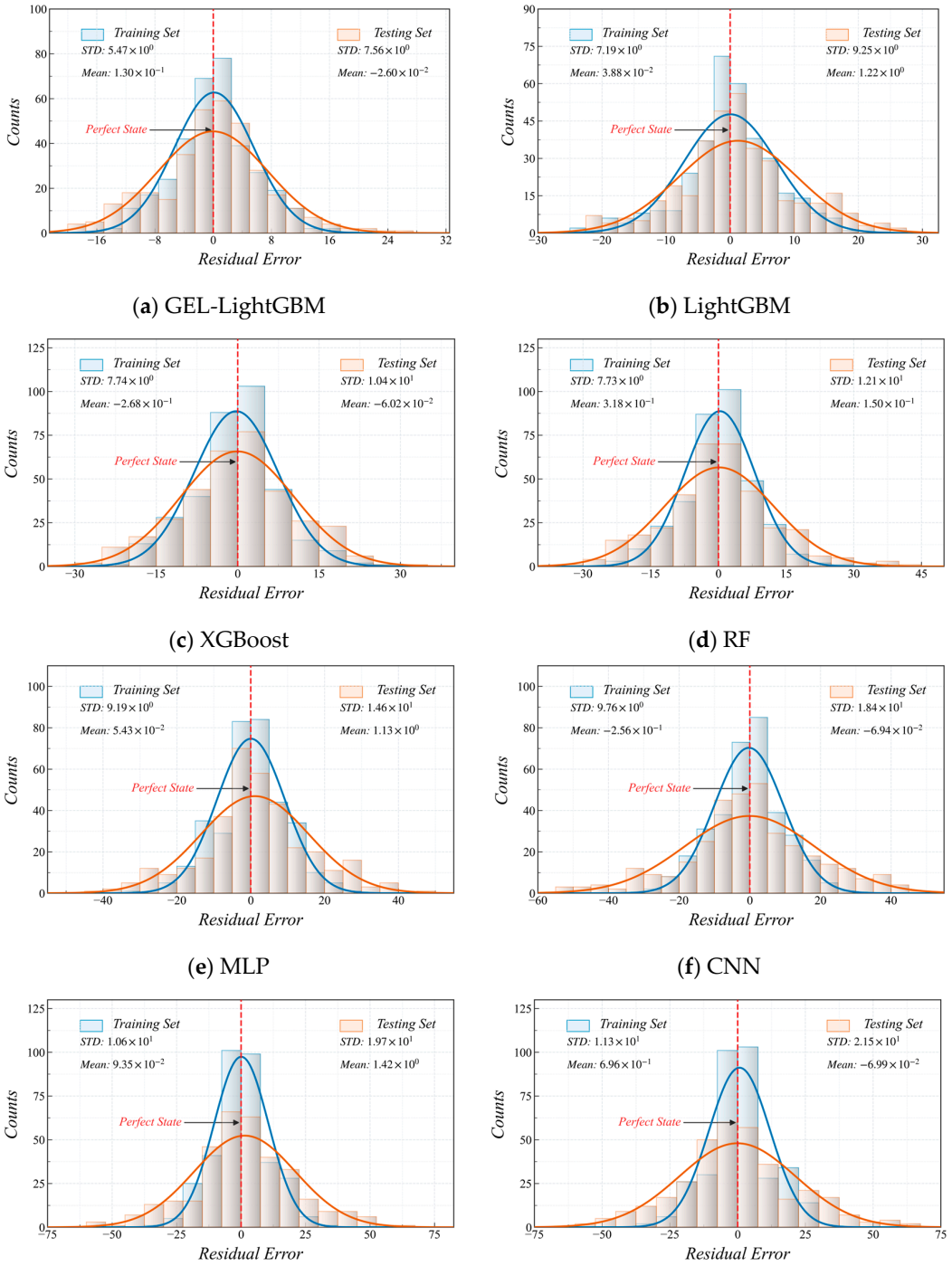


Figure 5. Histograms of residual distributions for prediction results.

4.2. Model Performance Evaluation

To further quantitatively evaluate the predictive capability of different models for rock uniaxial compressive strength (UCS), four evaluation metrics, namely R^2, RMSE, MAPE, and MAE, were calculated for both the training and testing sets. The results are listed in Table 2 and visualized using polar bar charts, as shown in Figure 6. Among these metrics, R^2 is used to characterize the explanatory ability of the model for variations in UCS, with values closer to 1 indicating better fitting performance. RMSE, MAPE, and MAE are used to measure prediction errors, where smaller values indicate higher prediction accuracy and lower error fluctuation. By examining the performance on both the training and testing sets, the fitting ability, generalization performance, and prediction stability of each model can be evaluated more comprehensively.

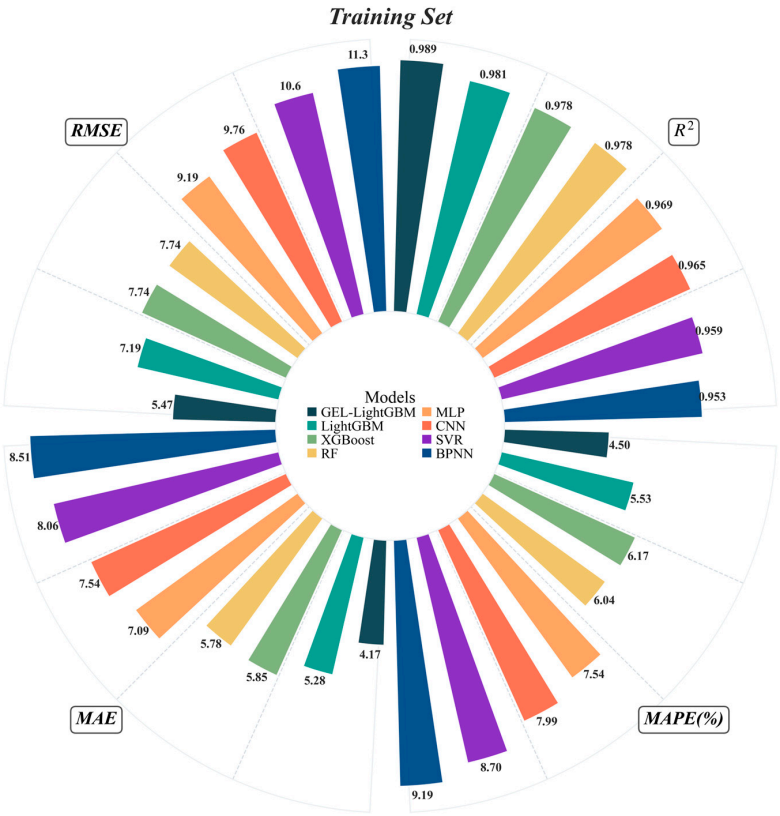
Table 2. Evaluation of model prediction performance.

Dataset	Model	R^2	RMSE	MAPE (%)	MAE
Training Set	GEL-LightGBM	0.9890	5.4732	4.4978	4.1697
	LightGBM	0.9810	7.1932	5.5309	5.2768
	XGBoost	0.9780	7.7403	6.1689	5.8542
	RF	0.9780	7.7403	6.0432	5.7804
	MLP	0.9690	9.1881	7.5446	7.0879
	CNN	0.9650	9.7629	7.9874	7.5390
	SVR	0.9590	10.5666	8.7018	8.0585
	BPNN	0.9530	11.3134	9.1882	8.5115
Testing Set	GEL-LightGBM	0.9790	7.5623	6.0490	5.6705
	LightGBM	0.9680	9.3351	7.2692	6.9725
	XGBoost	0.9600	10.4370	8.8775	8.1091
	RF	0.9460	12.1266	9.4184	9.0591
	MLP	0.9210	14.6675	11.6206	10.8832
	CNN	0.8760	18.3762	14.3993	13.7729
	SVR	0.8573	19.7131	16.5948	14.9169
	BPNN	0.8310	21.4530	17.3397	16.3338

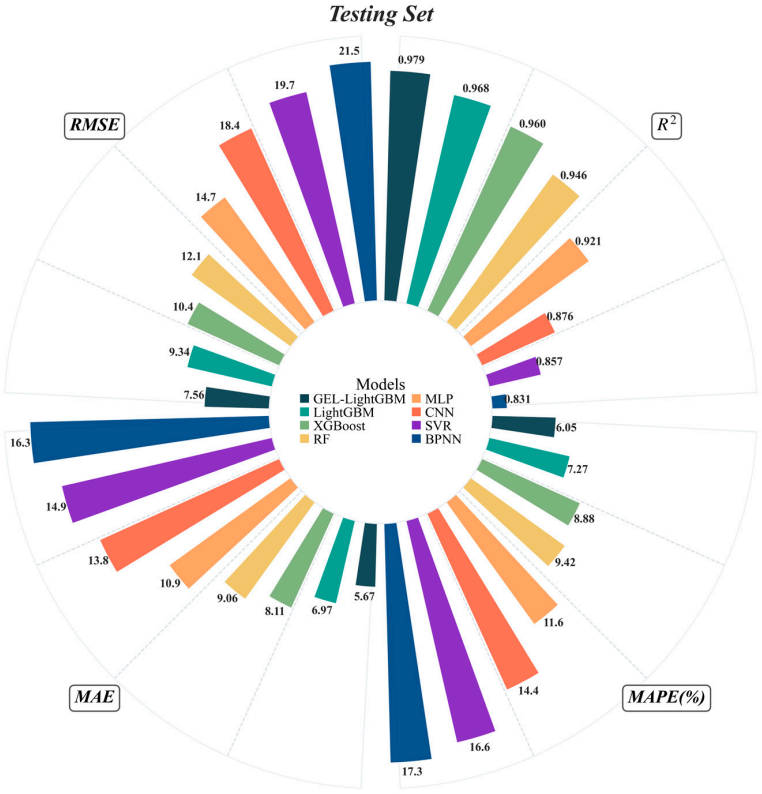
As shown in Table 2 and Figure 6, GEL-LightGBM achieves the best performance on both the training and testing sets. On the training set, GEL-LightGBM obtains an R^2 of 0.9890, with RMSE, MAPE, and MAE values of only 5.4732 MPa, 4.4978%, and 4.1697 MPa, respectively. On the testing set, its R^2 remains as high as 0.9790, while the RMSE, MAPE, and MAE are 7.5623 MPa, 6.0490%, and 5.6705 MPa, respectively. These results indicate that the proposed model not only has strong fitting ability on the training data, but also maintains good generalization accuracy on unseen samples.

Compared with conventional LightGBM, GEL-LightGBM reduces the testing RMSE, MAPE, and MAE by approximately 19.0%, 16.8%, and 18.7%, respectively. This demonstrates that, after introducing empirical-law priors, sample-adaptive gated fusion, and residual correction, the model can more effectively use rock mechanics knowledge to constrain the data-driven learning process. In comparison, ensemble learning models such as LightGBM, XGBoost, and RF generally outperform MLP, CNN, SVR, and BPNN, indicating that tree-based ensemble models are more suitable for UCS prediction problems involving small samples, multiple variables, and strong nonlinearity. However, these purely data-driven models lack explicit physical priors, and their testing errors remain higher than those of GEL-LightGBM. The neural-network-based models and SVR show relatively large testing errors. In particular, BPNN has a testing R^2 of only 0.8310 and an RMSE of 21.4530 MPa, suggesting that complex parametric models or kernel-based models are more susceptible to data distribution and parameter sensitivity when the sample size is limited, and lithological differences are significant, leading to insufficient generalization stability. The polar bar charts also visually reflect this trend: GEL-LightGBM corresponds to the

highest bar in the R^2 region and the lowest bars in the RMSE, MAPE, and MAE regions, forming the best overall performance combination.



(**a**) Training Set



(**b**) Testing Set

Figure 6. Polar bar charts of model performance.

In summary, GEL-LightGBM outperforms the comparison models across all four evaluation metrics, verifying its effectiveness and reliability in the UCS prediction task. Its performance advantage does not simply arise from increased model complexity, but from the synergistic effect among the empirical-law expert layer, gated weight allocation, and residual corrector. The empirical-law module provides prior constraints with rock mechanics significance, the gating mechanism enhances the model's adaptive selection of empirical relationships suitable for different samples, and residual correction further compensates for the nonlinear deviation between empirical predictions and measured strength. Therefore, GEL-LightGBM can improve model generalization while maintaining high prediction accuracy, providing a more engineering-applicable modeling scheme for rapid UCS prediction under complex lithological conditions.

4.3. Ablation Experiment

To verify the effectiveness of each component of GEL-LightGBM, ablation experiments were conducted in this study. Conventional LightGBM was set as the purely data-driven baseline. GEL-LightGBM (no-empirical) was constructed by removing the empirical-law expert module to examine the contribution of embedding classical empirical equations to prediction performance. GEL-LightGBM (no-gate) was constructed by replacing the sample-adaptive gated fusion with fixed empirical-law fusion to test the necessity of the gating mechanism. Finally, the complete GEL-LightGBM model was used for comparison. All models adopted the same data partition, input variables, and evaluation metrics.

As shown in Table 3, the complete GEL-LightGBM achieves the best performance on the testing set, with R^2, RMSE, MAPE, and MAE values of 0.9790, 7.5623, 6.0490%, and 5.6705, respectively. Compared with conventional LightGBM, GEL-LightGBM reduces the RMSE from 9.3351 to 7.5623 and the MAE from 6.9725 to 5.6705, indicating that the joint design of empirical-law embedding, gated fusion, and residual correction can effectively improve the generalization ability of the model.

Table 3. Results of the ablation experiment.

Model	R^2	RMSE	MAPE (%)	MAE
GEL-LightGBM	0.9790	7.5623	6.0490	5.6705
GEL-LightGBM (no-empirical)	0.9713	8.9314	7.0011	6.5432
GEL-LightGBM (no-gate)	0.9756	8.0172	6.5139	6.0147
LightGBM	0.9680	9.3351	7.2692	6.9725

The R^2 of GEL-LightGBM (no-empirical) is 0.9713, which is lower than that of the complete model. This indicates that after removing the empirical equations, the model still retains a certain nonlinear fitting capability but lacks the physical priors provided by point-load strength, the wave velocity-hardness comprehensive relationship, and the porosity-degradation law, resulting in increased prediction errors. GEL-LightGBM (no-gate) obtains an R^2 of 0.9756 and an RMSE of 8.0172, performing better than the no-empirical variant but worse than the complete model. This suggests that the empirical laws themselves can provide effective strength constraints, but fixed fusion cannot fully reflect the differences in the applicability of empirical equations among different rock samples. The complete model dynamically adjusts the contributions of different empirical laws through the sample-adaptive gating mechanism, allowing the coupling relationships among porosity, rebound hardness, wave velocity, and point-load strength to be represented more reasonably.

In summary, the ablation experiment shows that empirical-law embedding is the main source of performance improvement, while the gated fusion mechanism further enhances

the sample adaptability of the empirical priors. The complete GEL-LightGBM achieves a better balance between data-driven learning and empirical laws of rock strength.

4.4. Interpretability Analysis

To further reveal the internal basis of the model for predicting uniaxial compressive strength (UCS), the SHAP method was used for interpretability analysis [35]. Based on the concept of Shapley values, SHAP decomposes the model prediction into the contribution values of individual input variables, thereby quantitatively characterizing the positive or negative effects of different features on UCS prediction. A SHAP value greater than 0 indicates that the feature increases the predicted UCS, whereas a SHAP value less than 0 indicates that the feature decreases the predicted value. Compared with simple correlation analysis, SHAP can simultaneously characterize feature importance, nonlinear responses, and interactions among variables, making it more suitable for explaining the multi-factor coupling effects in rock strength prediction.

Figures 7 and 8 show the SHAP beeswarm plot and feature importance results, respectively. As shown in Figure 7, porosity (n) has the widest SHAP value distribution, and high-porosity samples are mainly located in the negative SHAP region, indicating that an increase in n significantly decreases the predicted UCS. In contrast, low-porosity samples correspond to clear positive contributions. This is consistent with rock mechanics mechanisms: pores and microcracks weaken the internal cemented structure of rock, reduce the effective load-bearing area, and induce stress concentration and crack propagation during compression, thereby decreasing compressive strength. High-value samples of SHR and V_p are mainly distributed in the positive SHAP region, indicating that higher rebound hardness and P-wave velocity increase the predicted UCS. SHR reflects rock surface hardness and local compressive resistance, whereas V_p reflects the internal compactness and integrity of rock. Higher values of these two variables generally indicate a more intact rock structure and better cementation.

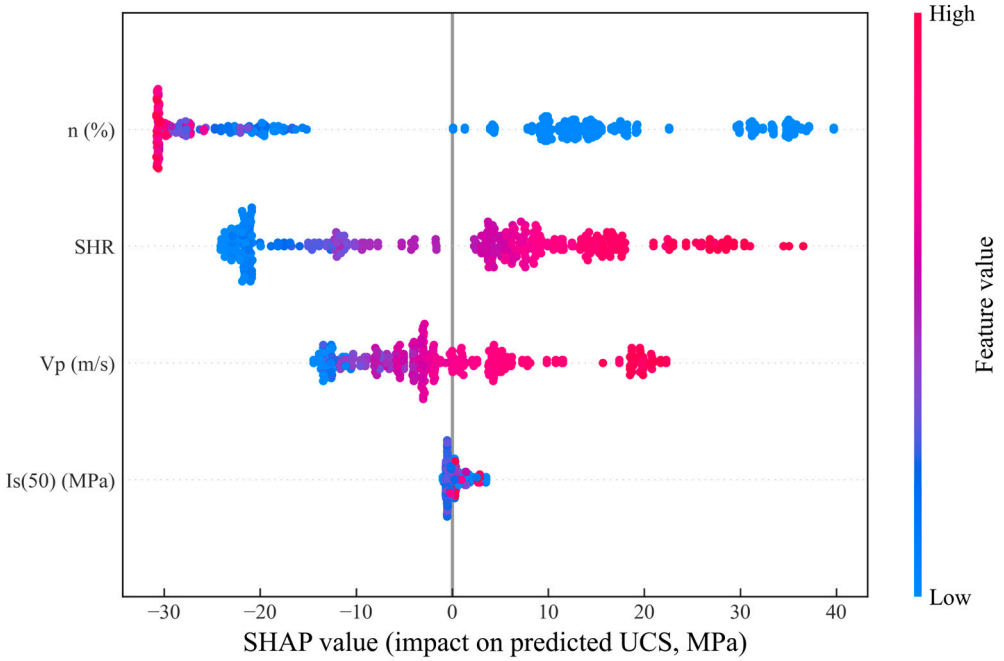


Figure 7. SHAP beeswarm plot.

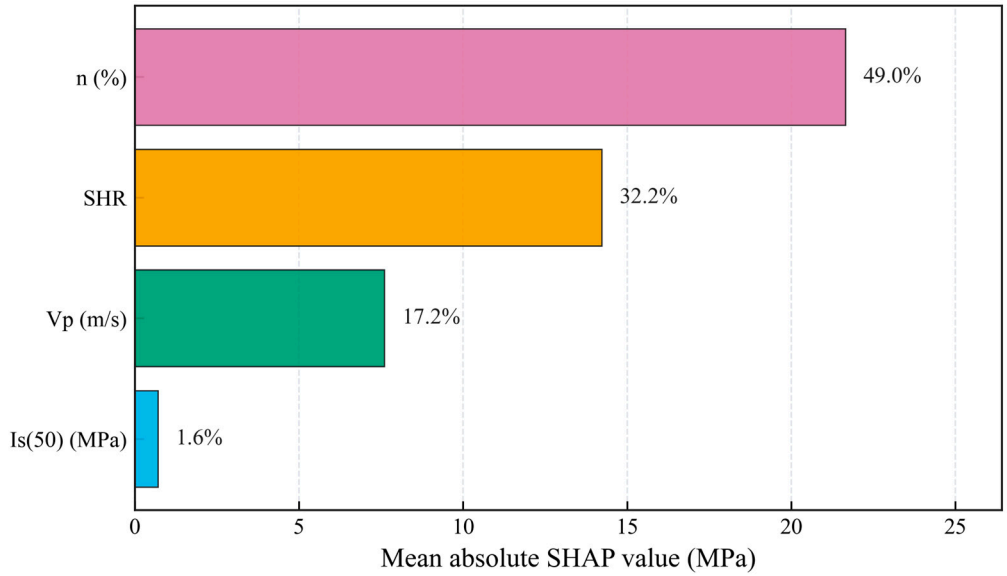


Figure 8. Feature importance results.

Figure 8 further shows that n has the highest mean absolute SHAP contribution, accounting for 49.0%, making it the dominant variable in UCS prediction. The contributions of SHR and V_p are 32.2% and 17.2%, respectively, indicating that hardness and wave velocity provide important supplementary information for strength prediction. The contribution of $I_s(50)$ is only 1.6%. This does not mean that $I_s(50)$ lacks engineering significance, but rather indicates that its marginal explanatory contribution is relatively limited under the current multivariable and empirical-law fusion framework.

Figures 9 and 10 further reveal the interaction effects among the variables. The pairwise interaction scatter plots in Figure 9 show clear differences in the pairwise SHAP interaction values corresponding to different variable combinations. Among them, the average interaction strength between n and SHR is the highest, reaching 2.31 MPa, followed by that between n and V_p, with an average interaction strength of 0.84 MPa. The interaction strength between SHR and V_p is 0.68 MPa.

The interaction heatmap in Figure 10 further confirms this pattern. The main effects of n, SHR, and V_p are significantly stronger than that of $I_s(50)$. Among them, n has the strongest main effect, reaching 23.60 MPa, while the main effects of SHR and V_p are 15.78 MPa and 7.86 MPa, respectively. The off-diagonal results indicate that the most significant coupling effect occurs between n and SHR. This suggests that pore structure not only directly controls UCS but also regulates the contributions of rebound hardness and wave velocity to strength. When rock porosity is low, higher SHR and V_p more clearly reflect the high load-bearing capacity of intact and dense rocks. However, when pores and fractures are more developed, even relatively high local hardness or wave velocity may have a weakened positive effect on the overall compressive strength. This result indicates that the model captures the coupled control mechanism among pore defects, surface hardness, and internal integrity.

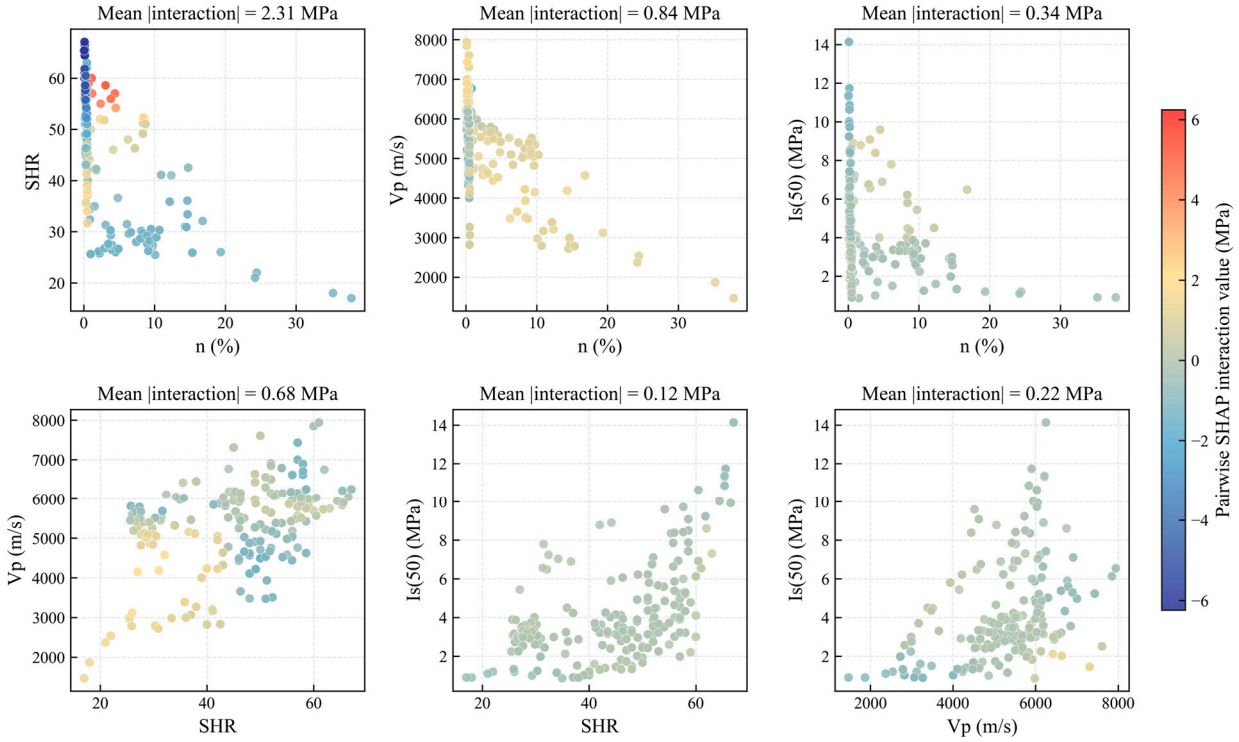


Figure 9. Feature interaction scatter plots.

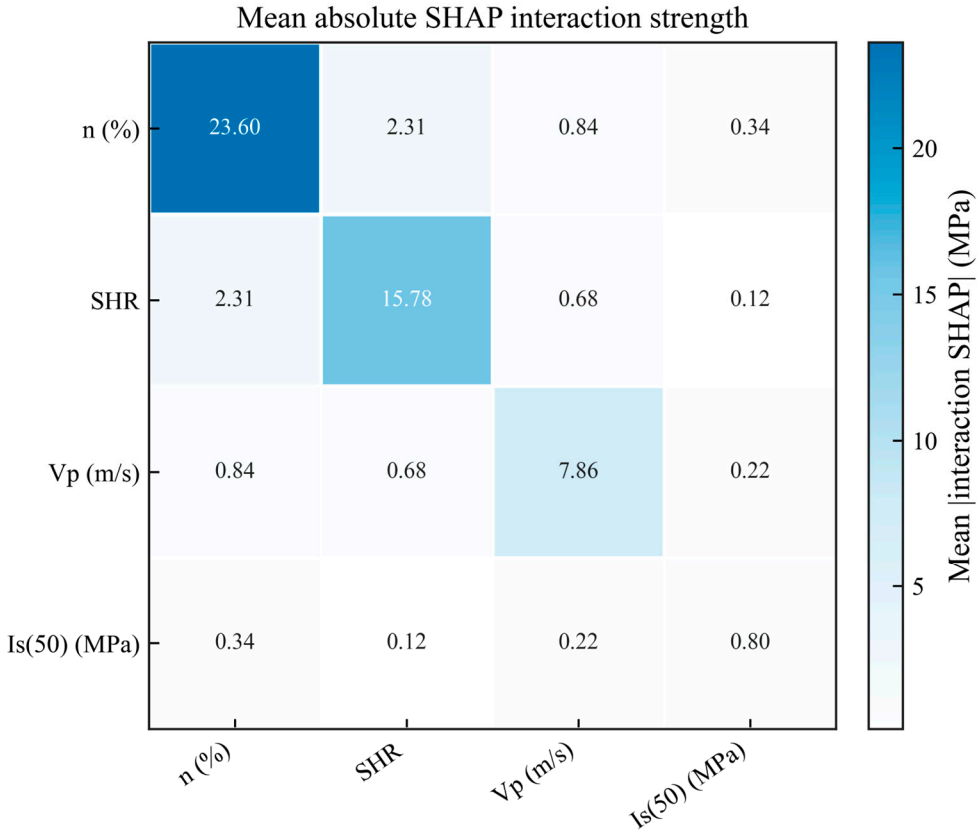


Figure 10. Heatmap of feature interaction effects.

Figure 11 presents the SHAP dependence relationships of each feature, which are used to analyze the nonlinear influence of changes in variable values on their contributions to UCS prediction. The SHAP value of porosity (n) decreases rapidly as n increases and remains clearly negative in the higher-porosity range, indicating that the weakening effect

of porosity on UCS has a significant nonlinear characteristic. The SHAP value of SHR generally increases with increasing SHR and shows a stronger positive contribution in the higher-SHR range, suggesting that increased rock surface hardness enhances the predicted UCS. The dependence curve of V_p also shows a clear upward trend. In particular, in the higher wave velocity range, the SHAP value rapidly becomes positive and continues to increase, indicating that denser internal structures and fewer fractures correspond to higher rock compressive strength. In contrast, the SHAP values of $I_s(50)$ show relatively small overall fluctuations, with most samples concentrated near zero. This indicates that its influence on UCS is relatively weak and is more reflected as an auxiliary correction effect. The interaction variables represented by color in Figure 11 also indicate certain coupling relationships between n and SHR and between n and V_p, which is consistent with the interaction analysis results in Figures 9 and 10.

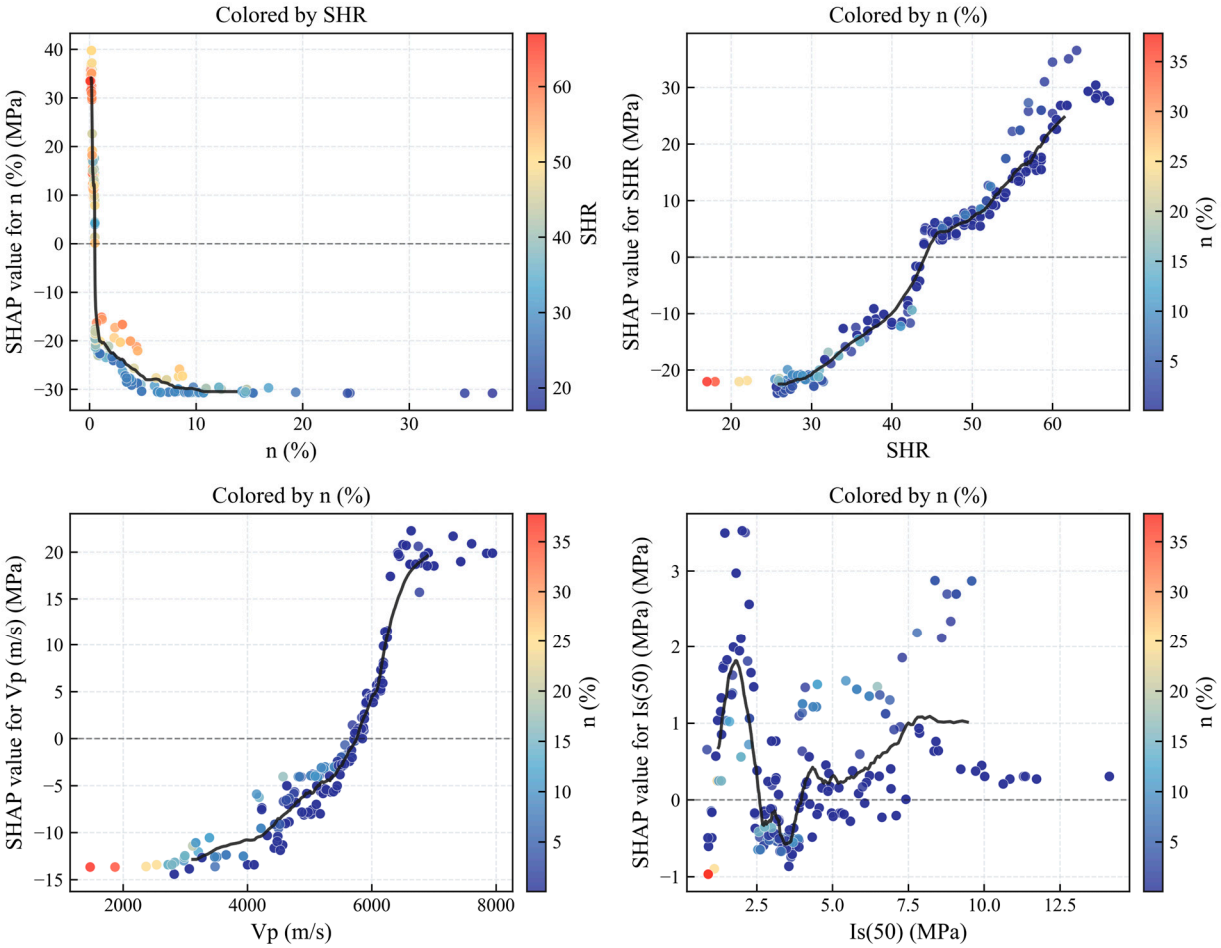


Figure 11. SHAP dependence plots.

In summary, the SHAP analysis indicates that porosity is the primary factor controlling UCS prediction, while SHR and V_p provide key supplementary information from the perspectives of surface hardness and internal integrity, respectively. In contrast, the independent contribution of $I_s(50)$ is relatively weak. The feature contribution directions, importance ranking, and interaction relationships identified by the model are consistent with rock mechanics mechanisms related to pore defects, cementation state, wave propagation, and compressive failure. This demonstrates that the established prediction model not only has good data-fitting capability but also possesses reasonable physical interpretability.

5. Discussion

5.1. Generalization Advantage Under Empirical Prior Constraints

The performance advantage of GEL-LightGBM indicates that, under small-sample, multi-lithology, and strongly nonlinear conditions, explicitly introducing empirical laws of rock strength is more beneficial for improving the stability of UCS prediction than relying solely on data-driven fitting. Although conventional LightGBM, XGBoost, and RF can characterize nonlinear relationships among variables, their learning processes are mainly controlled by the distribution of training samples. When the sample sources are complex and lithological differences are significant, purely data-driven models may form biases toward local data patterns, leading to increased testing errors.

In contrast, GEL-LightGBM organizes the point-load strength, multi-index comprehensive strength, and porosity-degradation relationships as empirical-law experts, allowing the model to be constrained by rock mechanics priors during fitting. The testing results show that GEL-LightGBM achieves an R^2 of 0.9790, with RMSE, MAPE, and MAE values of 7.5623 MPa, 6.0490%, and 5.6705 MPa, respectively, which are approximately 19.0%, 16.8%, and 18.7% lower than those of conventional LightGBM. The ablation experiment also shows that removing either the empirical-law experts or the gating mechanism leads to performance degradation, indicating that the improvement in prediction accuracy arises from the synergy among empirical priors, sample-adaptive fusion, and residual correction.

Therefore, the core value of GEL-LightGBM does not lie in simply stacking more model structures, but in transforming the physical constraints of empirical equations into learnable and adjustable model components. The empirical-law layer provides prior boundaries for strength variation, the gating mechanism regulates the applicability of different empirical relationships to specific samples, and the residual correction layer further compensates for the nonlinear deviation between empirical predictions and measured UCS. Together, these three components improve the generalization ability of the model for heterogeneous rock data.

5.2. Hardness-Wave Velocity Synergistic Response Regulated by Pore Defects

The SHAP analysis shows that porosity is the dominant variable affecting UCS prediction, with a mean absolute contribution of 49.0%, higher than the contributions of SHR and V_p, which are 32.2% and 17.2%, respectively. This result indicates that pore structure is not merely a single input variable but also a key defect descriptor regulating the load-bearing capacity of rock. Pores and microcracks weaken the continuity of the rock skeleton, reduce the effective load-bearing area, and promote stress concentration and crack propagation during compression, thereby causing UCS to decrease with increasing porosity.

SHR and V_p reflect rock surface hardness and internal integrity, respectively. A higher SHR usually corresponds to stronger local compressive resistance, while a higher V_p indicates a denser internal structure and fewer fractures. Their positive contributions to UCS indicate that surface hardness and internal integrity are both important supplementary pieces of information for rock strength prediction. However, these contributions do not occur independently, but are regulated by the state of the pore structure.

The feature interaction results further support this interpretation. The average interaction strength between porosity and SHR is the highest, reaching 2.31 MPa, while the interaction strength between porosity and V_p is 0.84 MPa, both higher than that between SHR and V_p. This indicates that, when porosity is low, high SHR and high V_p more clearly reflect the high load-bearing capacity of intact and dense rocks. When pores and fractures are well developed, however, even relatively high local hardness or wave velocity may have a weakened positive effect on the overall UCS. Therefore, UCS is controlled by the

coupling of pore defects, surface hardness, and internal integrity, rather than being linearly determined by a single indicator.

5.3. Applicability Boundaries and Extension Directions Under Multi-Source Heterogeneous Data

Although GEL-LightGBM achieves high prediction accuracy on the current database, it should not be regarded as a universal UCS prediction model for all rock types and engineering conditions. The 344 samples used in this study were collected from different publications and cover multiple common lithologies, including granite, limestone, sandstone, tuff, and travertine, rather than rocks from a single specific area. This multi-source composition improves the model's applicability to heterogeneous rock datasets, but it may also introduce inconsistencies in testing standards, specimen preparation, testing equipment, measurement procedures, lithological coverage, and data reporting formats. These source-related differences may increase the dispersion of the relationships between indirect indices and UCS, and may affect model performance when the trained model is transferred to independent engineering scenarios. Therefore, the model is expected to be more reliable for rock samples whose physical indices, lithological characteristics, and testing conditions fall within or close to the range represented by the current database.

In addition, the current model only uses porosity, SHR, V_p, and $I_s(50)$ as input variables, and does not explicitly consider factors such as mineral composition, grain structure, cementation type, water content, weathering degree, and microcrack geometry. These factors also influence UCS and may have different effects across lithologies. Therefore, when the rocks to be predicted differ substantially from the training data in lithological structure or testing conditions, uncertainty may still exist in the model predictions.

Future research should place greater emphasis on independent external validation. A larger and more standardized UCS database should be developed, and external datasets collected from different regions, laboratories, lithologies, or engineering projects should be used to test the transferability of GEL-LightGBM beyond the present data sources. In addition, lithology categories, mineralogical parameters, density, water content, and microstructural indices can be introduced to better represent the factors controlling rock strength. Uncertainty quantification, lithology-grouped modeling, and transfer learning may further improve the reliability of the model for small-sample lithologies and complex engineering conditions. In this way, GEL-LightGBM can be further developed from a high-accuracy prediction model on the current database into a more robust UCS evaluation method with broader engineering applicability.

6. Conclusions

This study proposed a Gated Empirical-Law LightGBM model (GEL-LightGBM) for rapid UCS prediction under multi-source rock sample conditions. The model embeds point-load strength, multi-index comprehensive strength, and porosity-degradation relationships as empirical-law experts, and achieves synergistic modeling of rock mechanics priors and data-driven learning through sample-adaptive gated fusion and controlled residual correction. Based on 344 publicly available rock mechanics test samples, the following conclusions are drawn:

(1) GEL-LightGBM shows the best overall performance in UCS prediction. On the testing set, the model achieves R^2, RMSE, MAPE, and MAE values of 0.9790, 7.5623 MPa, 6.0490%, and 5.6705 MPa, respectively, outperform comparison models including LightGBM, XGBoost, RF, MLP, CNN, SVR, and BPNN. Compared with conventional LightGBM, its testing RMSE, MAPE, and MAE are reduced by approximately 19.0%, 16.8%, and 18.7%, respectively, indicating that empirical-law constraints help improve prediction accuracy and generalization stability for multi-lithology data.

(2) The ablation experiment verifies the effectiveness of each model component. After removing the empirical-law experts, the model error increases noticeably. When sample-adaptive gating is replaced by fixed fusion, the prediction performance also declines. These results indicate that empirical-law embedding is an important source of performance improvement, while the gating mechanism can further identify the applicability of different empirical equations to different rock samples. Controlled residual correction compensates for the nonlinear deviation between empirical priors and measured UCS, preventing the model from relying entirely on either empirical equations or purely data-driven fitting.

(3) The SHAP analysis shows that porosity is the dominant variable in UCS prediction, with a mean absolute contribution of 49.0%. The contributions of SHR and V_p are 32.2% and 17.2%, respectively, while the independent contribution of $I_s(50)$ is 1.6%. The feature contribution directions are consistent with rock mechanics understanding: increasing porosity weakens the load-bearing capacity of rock, whereas higher SHR and V_p reflect stronger surface hardness and better internal integrity, respectively.

(4) The feature interaction results further reveal that UCS is controlled by multi-factor coupling. The average interaction strength between porosity and SHR is the highest, reaching 2.31 MPa, while the interaction strength between porosity and V_p is 0.84 MPa. This indicates that pore structure not only directly affects UCS but also regulates the contributions of hardness and wave velocity information to strength prediction. Therefore, GEL-LightGBM not only improves prediction accuracy but also provides a physical interpretation consistent with pore defects, surface hardness, and internal integrity.

Despite these promising results, the applicability of GEL-LightGBM is still influenced by the scale and composition of the current database. The 344 samples used in this study were compiled from published studies, and variations in testing standards, specimen preparation, lithological coverage, and data recording may introduce uncertainty when the model is applied to independent engineering cases. In addition, the current input system is limited to porosity, SHR, V_p, and $I_s(50)$, whereas mineral composition, water content, weathering degree, grain structure, and microcrack characteristics are not explicitly included.

Future research should therefore focus on building larger and more standardized UCS databases, validating the model on independent external datasets, and incorporating additional lithological and microstructural descriptors. Combining GEL-LightGBM with uncertainty quantification, lithology-grouped modeling, or transfer learning may further improve its reliability for complex engineering applications. Overall, the proposed framework provides a physically informed and interpretable modeling route for rapid UCS prediction, while leaving clear directions for broader validation and practical extension.

Author Contributions: Conceptualization, J.P. and H.L.; methodology, H.L.; software, J.P.; validation, Y.L. and Y.H.; formal analysis, Y.H.; investigation, J.P.; resources, H.L.; data curation, Y.L.; writing—original draft preparation, J.P.; writing—review and editing, H.L. and Y.L.; visualization, J.P. and Y.H.; supervision, H.L. and Y.L.; project administration, Y.L.; funding acquisition, Y.L. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The original contributions presented in the study are included in the article; further inquiries can be directed to the corresponding author.

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

Abbreviations

The following abbreviations are used in this manuscript:

UCS	Uniaxial Compressive Strength
GEL-LightGBM	Gated Empirical-Law LightGBM
SHR	Schmidt Rebound Hardness
RF	Random Forest
MLP	Multilayer Perceptron
CNN	Convolutional Neural Network
SVR	Support Vector Regression
BPNN	Back-Propagation Neural Network
SHAP	Shapley Additive Explanations
KNN	K-Nearest Neighbor

References

1. Xie, W.-Q.; Liu, X.-L.; Zhang, X.-P.; Liu, Q.-S.; Wang, E.-Z. A Review of Test Methods for Uniaxial Compressive Strength of Rocks: Theory, Apparatus and Data Processing. *J. Rock Mech. Geotech. Eng.* **2025**, *17*, 1889–1905.
2. Sabri, M.S.; Verma, A.K.; Singh, T. Direct and Indirect Methods for Uniaxial Compressive Strength Estimation in Various Geo-Environments: A Review. *J. Rock Mech. Geotech. Eng.* **2025**, *17*, 8241–8275. [CrossRef]
3. Du, K.; Li, X.; Su, R.; Tao, M.; Lv, S.; Luo, J.; Zhou, J. Shape Ratio Effects on the Mechanical Characteristics of Rectangular Prism Rocks and Isolated Pillars under Uniaxial Compression. *Int. J. Min. Sci. Technol.* **2022**, *32*, 347–362. [CrossRef]
4. González-Fernández, M.A.; Estévez-Ventosa, X.; Alejano, L.R.; Masoumi, H. Size-Dependent Behaviour of Hard Rock under Triaxial Loading. *Rock Mech. Rock Eng.* **2023**, *56*, 6009–6025. [CrossRef]
5. Kong, F.; Xue, Y.; Shang, J.; Zhu, C.; Han, M.; Qu, Z.; Wang, B.; Yang, K. Predicting Uniaxial Compressive Strength of Building Stone Based on Index Tests: Correlations, Validity, Reliability, and Unification. *Constr. Build. Mater.* **2024**, *438*, 137227. [CrossRef]
6. Lobarinhas, R.; Dionísio, A.; Paneiro, G.; e Silva, M.C. Evaluating Leeb Hardness as a Reliable UCS Predictor for Stone Masonry Exposed to High Temperatures: The Influence of Porosity across Diverse Carbonate Lithologies. *J. Build. Eng.* **2025**, *111*, 113079. [CrossRef]
7. Jamshidi, A. A Comparative Study of Point Load Index Test Procedures in Predicting the Uniaxial Compressive Strength of Sandstones. *Rock Mech. Rock Eng.* **2022**, *55*, 4507. [CrossRef]
8. Skentou, A.D.; Bardhan, A.; Mamou, A.; Lemonis, M.E.; Kumar, G.; Samui, P.; Armaghani, D.J.; Asteris, P.G. Closed-Form Equation for Estimating Unconfined Compressive Strength of Granite from Three Non-Destructive Tests Using Soft Computing Models. *Rock Mech. Rock Eng.* **2023**, *56*, 487–514.
9. Setayeshirad, M.R.; Uromeie, A.; Nikudel, M.R. Modeling of the Uniaxial Compressive Strength of Carbonate Rocks Using Statistical Analysis and Artificial Neural Network. *Rock Mech. Rock Eng.* **2025**, *58*, 6025–6042. [CrossRef]
10. Hussain, J.; Ali, N.; Fu, X.; Chen, J.; Khan, N.A.; Hussain, S.; Iqbal, S.M. Prediction of Uniaxial Compressive Strength from Petrographic Features of Kirana Dolerites Using a Hybrid Group Method of Data Handling-Levenberg–Marquardt (GMDH-LM) Model. *Rock Mech. Rock Eng.* **2026**, *59*, 1617–1645.
11. Zhou, S.; Zhang, Z.-X.; Luo, X.; Huang, Y.; Yu, Z.; Yang, X. Predicting Dynamic Compressive Strength of Frozen-Thawed Rocks by Characteristic Impedance and Data-Driven Methods. *J. Rock Mech. Geotech. Eng.* **2024**, *16*, 2591–2606.
12. Xie, N.; Luo, X.; Chen, J.; Yan, S. Machine Learning Models for Prediction of Sandstone Compressive Strength Under Hydro-Chemical Corrosion. *Rock Mech. Rock Eng.* **2025**, *58*, 13349–13367. [CrossRef]
13. Ye, T.; Chen, Z.; Zhang, C.; Li, C.; Wei, W.; Dong, J. A Novel XRF and Machine Learning-Based Method for Predicting Weathered Rock Strength. *Int. J. Rock Mech. Min. Sci.* **2025**, *191*, 106130. [CrossRef]
14. Liu, X.; Xu, D.; Li, S.; Duan, S.; Xu, H.; Jiang, Q.; Qiu, S. Estimating the Mechanical Properties of Rocks and Rock Masses Based on Mineral Micromechanics Testing. *Rock Mech. Rock Eng.* **2024**, *57*, 5267–5278. [CrossRef]
15. Piao, S.; Huang, S.; Wei, Y.; Tan, J.; Ma, B. A Novel Hybrid Bayesian-Group-Based Machine Learning (HB-GML) Method for Predicting Uniaxial Compressive Strength (UCS) of Rock. *Rock Mech. Rock Eng.* **2025**, *58*, 5817–5843.
16. Liu, Z.; Li, D.; Liu, Y.; Yang, B.; Zhang, Z.-X. Prediction of Uniaxial Compressive Strength of Rock Based on Lithology Using Stacking Models. *Rock Mech. Bull.* **2023**, *2*, 100081. [CrossRef]
17. Khatti, J.; Grover, K.S. Prediction of Uniaxial Strength of Rocks Using Relevance Vector Machine Improved with Dual Kernels and Metaheuristic Algorithms. *Rock Mech. Rock Eng.* **2024**, *57*, 6227–6258. [CrossRef]
18. Zhao, T.; Song, C.; Lu, S.; Xu, L. Prediction of Uniaxial Compressive Strength Using Fully Bayesian Gaussian Process Regression (fB-GPR) with Model Class Selection. *Rock Mech. Rock Eng.* **2022**, *55*, 6301–6319.

19. Davoodi, S.; Mehrad, M.; Wood, D.A.; Rukavishnikov, V.S.; Bajolvand, M. Predicting Uniaxial Compressive Strength from Drilling Variables Aided by Hybrid Machine Learning. *Int. J. Rock Mech. Min. Sci.* **2023**, *170*, 105546. [CrossRef]
20. Zhao, R.; Shi, S.; Yao, R.; Yang, S. Application of Relationship Model for the Measurement While Drilling Data to Predict Rock Uniaxial Compressive Strength for Tunneling. *Rock Mech. Rock Eng.* **2024**, *57*, 7187–7203. [CrossRef]
21. Xu, P.; Ji, X.; Li, M.; Lu, W. Small Data Machine Learning in Materials Science. *npj Comput. Mater.* **2023**, *9*, 42. [CrossRef]
22. Li, H.; Hosseini, S.; Gordan, B.; Zhou, J.; Ullah, S. Dimensionless Machine Learning: Dimensional Analysis to Improve LSSVM and ANN Models and Predict Bearing Capacity of Circular Foundations. *Artif. Intell. Rev.* **2025**, *58*, 117. [CrossRef]
23. Mahmoodzadeh, A.; Mohammadi, M.; Hashim Ibrahim, H.; Nariman Abdulhamid, S.; Ghafoor Salim, S.; Farid Hama Ali, H.; Kamal Majeed, M. Artificial Intelligence Forecasting Models of Uniaxial Compressive Strength. *Transp. Geotech.* **2021**, *27*, 100499. [CrossRef]
24. Kılıç, A.; Teymen, A. Determination of Mechanical Properties of Rocks Using Simple Methods. *Bull. Eng. Geol. Environ.* **2008**, *67*, 237–244. [CrossRef]
25. Dehghan, S.; Sattari, G.; Chehreh Chelgani, S.; Aliabadi, M. Prediction of Uniaxial Compressive Strength and Modulus of Elasticity for Travertine Samples Using Regression and Artificial Neural Networks. *Min. Sci. Technol. China* **2010**, *20*, 41–46. [CrossRef]
26. Armaghani, D.J.; Tonnizam Mohamad, E.; Momeni, E.; Monjezi, M.; Sundaram Narayanasamy, M. Prediction of the Strength and Elasticity Modulus of Granite through an Expert Artificial Neural Network. *Arab. J. Geosci.* **2015**, *9*, 48. [CrossRef]
27. Aladejare, A.E. Evaluation of Empirical Estimation of Uniaxial Compressive Strength of Rock Using Measurements from Index and Physical Tests. *J. Rock Mech. Geotech. Eng.* **2020**, *12*, 256–268. [CrossRef]
28. Benesty, J.; Chen, J.; Huang, Y.; Cohen, I. Pearson Correlation Coefficient. In *Noise Reduction in Speech Processing*; Springer: Berlin/Heidelberg, Germany, 2009; pp. 1–4.
29. Pujianto, U.; Wibawa, A.P.; Akbar, M.I. Others K-Nearest Neighbor (k-NN) Based Missing Data Imputation. In *Proceedings of the 2019 5th International Conference on Science In Information Technology (ICSITech)*; IEEE: New York, NY, USA, 2019; pp. 83–88.
30. Huang, L.; Qin, J.; Zhou, Y.; Zhu, F.; Liu, L.; Shao, L. Normalization Techniques in Training Dnns: Methodology, Analysis and Application. *IEEE Trans. Pattern Anal. Mach. Intell.* **2023**, *45*, 10173–10196. [CrossRef] [PubMed]
31. 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*, 3149–3157.
32. Broch, E.; Franklin, J. The Point-Load Strength Test. *Int. J. Rock Mech. Min. Sci. Geomech. Abstr.* **1972**, *9*, 669–676. [CrossRef]
33. Franklin, J.A. Suggested Method for Determining Point Load Strength. *Int. J. Rock Mech. Min. Sci. Geomech. Abstr.* **1985**, *22*, 51–60. [CrossRef]
34. Çobanoğlu, İ.; Çelik, S.B. Estimation of Uniaxial Compressive Strength from Point Load Strength, Schmidt Hardness and P-Wave Velocity. *Bull. Eng. Geol. Environ.* **2008**, *67*, 491–498. [CrossRef]
35. Timilsina, M.S.; Sen, S.; Uprety, B.; Patel, V.B.; Sharma, P.; Sheth, P.N. Prediction of HHV of Fuel by Machine Learning Algorithm: Interpretability Analysis Using Shapley Additive Explanations (SHAP). *Fuel* **2024**, *357*, 129573.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.