Mining of Mineral Deposits
Volume 20 (2026), Issue 2, 89-100

https://doi.org/10.33271/mining20.02.089

Prediction of rock mechanical parameters using gamma ray log data with regression and machine learning approaches

Andyono Broto Santoso [1,2*], Supandi Sujatono [3], Muhammad Alfiza Farhan [1], Rizky Syaputra [1], Basuki Rahmad [2], Barlian Dwinagara [2]

[1] *Institut Teknologi Sains Bandung, Jawa Barat, Indonesia*
[2] *Universitas Pembangunan Nasional Veteran Yogyakarta, Yogyakarta, Indonesia*
[3] *Institut Teknologi Nasional Yogyakarta, Yogyakarta, Indonesia*

**Corresponding author: e-mail ab.santoso@itsb.ac.id*

Abstract

Purpose. This study aims to investigate the feasibility of using gamma ray (GR) wireline logs as a principal predictor for estimating key rock mechanical parameters, including internal friction angle (φ), cohesion (c), and uniaxial compressive strength (UCS), as a rapid and cost-effective alternative to conventional laboratory testing.

Methods. A dataset of 54 depth-matched records, comprising GR log values and laboratory-derived mechanical properties, was compiled from multiple boreholes in sedimentary formations. Predictive models were developed using linear regression, Random Forest (RF), Extreme Gradient Boosting (XGBoost), and Artificial Neural Networks (ANN). Data preprocessing included median imputation, feature scaling, and an 80/20 train-test split. Model performance was evaluated using the coefficient of determination (R^2), mean squared error (MSE), and root mean squared error (RMSE). SHapley Additive exPlanations (SHAP) values were used to improve model interpretability.

Findings. Correlation analysis revealed strong positive associations between GR and UCS ($r = 0.852$), cohesion ($r = 0.799$), and friction angle ($r = 0.785$). Among machine learning models, XGBoost delivered the best performance for cohesion ($R^2 = 0.6764$, RMSE = 5.09 kPa) and UCS ($R^2 = 0.4427$, RMSE = 145.42 MPa), while RF was optimal for friction angle ($R^2 = 0.4955$, RMSE = 2.35°). ANN consistently underperformed relative to tree-based models. Linear regression provided competitive baseline performance, indicating a predominantly linear relationship between GR logs and mechanical properties at wireline resolution.

Originality. This work systematically positions GR logs as the principal predictor of rock mechanical properties – an approach rarely adopted in prior research that predominantly relied on sonic or density logs. The study provides the first systematic empirical comparison of linear and nonlinear models using GR as the primary input feature, establishing practical regression equations directly applicable to field deployment.

Practical implications. The derived regression equations and validated modeling framework enable engineers to perform rapid geomechanical assessments using widely available GR log data, thereby reducing dependence on costly, time-consuming laboratory testing. This approach is particularly valuable during early exploration stages and in mining contexts where core sample availability is limited.

Keywords: *gamma ray log; rock mechanical parameters; linear regression; machine learning; geomechanics*

1. Introduction

Reliable knowledge of rock mechanical properties underpins nearly every decision in mining and geotechnical engineering. Parameters such as the internal friction angle (φ), cohesion (c), and uniaxial compressive strength (UCS) govern the stability of slopes, tunnels, open-pit excavations, and wellbores; when these values are poorly estimated, designs can become either unsafe or uneconomically conservative [1]. In practice, however, obtaining these properties through standard laboratory tests triaxial compression, uniaxial loading, and associated index procedures is time-consuming, destructive, and expensive. Worse still, high-quality core samples are not always available, particularly in the early stages of exploration or in deep-drilling campaigns where continuous coring is impractical [1]-[3]. This persistent tension between the need for accurate geomechanical data and the practical limitations of direct testing has driven a sustained search for indirect prediction methods.

Among the key parameters, uniaxial compressive strength (UCS) occupies a central position. It serves as a fundamental input for slope stability analysis, tunnel support design, and mine planning, making it one of the most widely referenced mechanical indices in rock engineering [4], [5]. The accurate determination of UCS is essential for assessing how rock

Received: 31 December 2025. Accepted: 1 June 2026. Available online: 30 June 2026

© 2026. A.B. Santoso et al.
Mining of Mineral Deposits. ISSN 2415-3443 (Online) | ISSN 2415-3435 (Print)
This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.

masses behave under loading, especially in underground and surface excavation projects where failure can have catastrophic consequences [6], [7]. Yet the direct measurement of UCS through laboratory testing of intact specimens remains constrained by the same limitations of cost, time, and sample availability that affect geomechanical testing more broadly [1], [3]. These constraints have long motivated the development of indirect estimation methods that leverage more accessible rock properties physical indices, petrographic characteristics, and geophysical data as proxies for strength [4], [5], [8].

Historically, the simplest and most widely adopted indirect approaches have relied on empirical correlations with index tests. Point load index, Schmidt hammer rebound, and porosity measurements, for example, have been used extensively to estimate UCS owing to their simplicity and field applicability [3], [4], [9]. Petrographic parameters such as grain size, mineralogical composition, and pore structure have also been recognized as key controls on rock strength, particularly in sedimentary formations [8], [10], [11]. Index-based estimation has been extended to volcanic and tuffaceous rocks as well, where welding degree and matrix composition influence the mechanical response [1], [12], [13]. Although these empirical models have proven useful in specific settings, they are inherently limited by their linear assumptions and typically lack generalizability across lithologies, given the non-linear and complex nature of rock behaviour [9], [14]. Statistical regression techniques were subsequently introduced to enable multi-parameter modelling with improved accuracy [5], [8]. Even so, traditional regression approaches struggle to capture high-dimensional interactions among input variables, a shortcoming that becomes especially pronounced in heterogeneous geological formations [15], [16].

The emergence of machine learning (ML) over the past two decades has offered a powerful alternative. Artificial neural networks (ANNs), support vector machines (SVMs), and ensemble tree-based models have consistently demonstrated superior performance over conventional methods by capturing non-linear relationships and handling large, multi-dimensional datasets [17]-[19]. For instance, convolutional neural networks (CNNs) have been applied to predict the UCS of sandy dolomite, achieving higher accuracy than classical regression analysis [20]. Hybrid approaches that combine optimization algorithms such as the firefly algorithm and Bayesian optimization with ML models have further enhanced predictive accuracy and model robustness [21]-[23]. Advanced architectures, including CNNs and stacking ensemble models, have also shown considerable promise for handling spatial heterogeneity and multi-lithology scenarios, suggesting that the ceiling for ML-based geomechanical prediction has not yet been reached [15], [16], [24].

The comparative performance of different ML algorithms is itself an active area of investigation, and the findings carry direct relevance for geomechanical applications. Broader methodological comparisons have shown that XGBoost and Random Forest exhibit distinct advantages depending on dataset characteristics and prediction targets [25], findings subsequently validated in geomechanical contexts by Hassan and Arman [26] and Abdelhedi et al. [27], who compared multiple ML techniques for predicting UCS of carbonate rocks. Li et al. [24] showed that ensemble intelligence methodologies outperform single-learner approaches in complex rock engineering environments, while Liu et al. [25] applied multiple tree-based ML methods to estimate rock strength with encouraging results. Beyond tree-based models, Mohamad et al. [28] introduced a particle swarm optimization-based ANN (PSO-ANN) for predicting UCS of soft rocks, highlighting the value of metaheuristic optimization in improving neural network performance, while Wei et al. [29] further demonstrated that machine learning approaches coupled with optimization can achieve competitive accuracy in predicting rock compressive strength under varying conditions.

Adaptive neuro-fuzzy inference systems (ANFIS) represent yet another paradigm, one that integrates data-driven flexibility with interpretable rule-based logic. The foundational work by Jang [30] and Jang and Sun [31] established ANFIS as a robust framework for modelling non-linear relationships under uncertainty, and subsequent applications have extended the approach to diverse geotechnical domains. Al-Hamed et al. [32], for example, predicted soil fractions using natural radionuclide concentrations, while Cabalar et al. [33] modelled the constitutive behaviour of sand mixtures under undrained loading, demonstrating the method's capacity to replicate complex material responses. In the specific context of rock strength prediction, Cao [34] provided a systematic comparison of hybrid ANFIS configurations, establishing performance benchmarks that remain a reference for this model class.

Despite these advances, several practical and conceptual challenges persist. Recent studies have highlighted the use of ML models such as Random Forest and Support Vector Regression for real-time UCS prediction using measurement-while-drilling (MWD) data, offering a potentially cost-effective alternative to laboratory testing [35]. However, the adoption of complex AI models raises legitimate concerns about the "black box" problem: the difficulty in establishing transparent relationships between predictors and rock strength makes interpretability a critical focus for field deployment [36]. Furthermore, some researchers have pointed to the fundamental limitations of purely data-driven approaches, arguing that rock strength models must integrate mechanical principles and physical laws rather than relying solely on pattern recognition [37]. To address these concerns, advanced frameworks including hybrid ANFIS architectures and deep learning are being explored for robust and efficient prediction in dynamic geological environments [34], [38].

Coal-bearing sedimentary sequences present specific geomechanical challenges that motivate the development of lithology-sensitive prediction frameworks. Liang et al. [39] systematically characterized rock mechanical properties in coal measures of the Longtan Formation, establishing empirical baselines for strength and deformability across interbedded strata. Mahmoodzadeh et al. [40] applied AI-based UCS forecasting models using diverse geological inputs, while Kumar et al. [41] demonstrated the applicability of machine learning for lithology prediction from well log data in coalfield settings. Muzamhindo and Ferentinou [42] developed a generic compressive strength prediction model applicable to multiple lithologies, addressing the persistent challenge of transferability across sedimentary sequences. In the Indonesian geological context directly relevant to the present study Sujatono [43] established empirical relationships between geophysical log responses and shear strength parameters in sedimentary rocks, underscoring the practical importance of reliable geomechanical estimates for mine planning.

Geophysical well logs offer a natural pathway for overcoming the limitations of discrete laboratory testing. Sonic, density, neutron, gamma-ray, and resistivity logs provide continuous, in-situ measurements along the entire borehole, making them ideal candidates for indirect UCS estimation [5], [44]. The pioneering work of McNally [45], [46] established the feasibility of estimating rock strength from sonic and neutron logs in coal measures, laying the foundation for modern log-based prediction models. Subsequent research expanded this approach to various lithologies, including ignimbrites, sandstones, and carbonate rocks [1], [6], [47]. More recently, the integration of machine learning with geophysical logs has enabled high-resolution, continuous profiling of UCS along boreholes, significantly enhancing subsurface characterization in both mining and petroleum contexts [48]-[50].

Recent advancements have broadened both the types of log data and the modelling strategies employed. Elemental capture spectroscopy (ECS) logs have been used to infer mineralogical composition and its influence on rock strength [51], while logging-while-drilling (LWD) and MWD data now enable real-time geomechanical evaluation during drilling operations [35], [52]. Balaguera et al. [53] demonstrated the applicability of ML to subsurface physical property prediction across heterogeneous rock types using multiple inputs, and Sun et al. [54] predicted the composite strength of coal-grout systems using ensemble methods, further broadening the scope of log-based mechanical property estimation. In deep mining settings, Zhao et al. [55] applied boosting-based ML combined with optimization algorithms for UCS prediction, achieving strong performance metrics on rock core data. Of particular relevance to this study, the Sujatono series of investigations [43], [56], [57] established empirical relations between geophysical log responses, shear strength parameters, and mineralogical controls in sedimentary rocks of the Indonesian geological context, providing a local reference framework directly applicable to the present investigation.

Within this landscape, gamma-ray (GR) logs emerge as the primary variable of interest in this study. GR logs provide a continuous, inexpensive record of lithological variability that is particularly sensitive to clay-rich intervals governing rock competence. The radioactive elements potassium, thorium, and uranium commonly associated with clay minerals directly influence the gamma-ray signal, making it a reliable proxy for identifying zones of mechanical weakness or strength across sedimentary sequences [41], [44]. Unlike laboratory testing, which is limited to discrete core samples, GR logs capture high-resolution vertical variations and enable the recognition of subtle lithological boundaries and heterogeneities that strongly impact rock mechanical behaviour [58]. This continuous coverage, combined with documented statistical correlations between GR and mechanical parameters such as UCS, cohesion, and internal friction angle, explains why gamma-ray can serve not only as a lithological indicator but also as an integrative predictor in geomechanical modelling [1], [43]. Importantly, its near-universal availability in almost every borehole makes it a practical, field-deployable feature that bridges the gap between geophysics and rock mechanics, offering a cost-effective pathway for predictive analytics in mining and geotechnical engineering.

Despite this significant progress, several important challenges remain. First, many existing models have been developed for specific lithologies and lack transferability across diverse geological settings [42], [53]. Second, while hybrid ML models show improved accuracy, their interpretability and generalization capability require further validation, especially when trained on limited datasets [26]. Third, there is a notable scarcity of studies that systematically compare the performance of different hybrid ANFIS configurations in UCS prediction, despite their theoretical advantages in handling uncertainty and non-linearity [34]. Finally, although geophysical logs are widely available, their full potential remains underutilized due to inadequate integration strategies between raw log responses, petrophysical transformations, and mechanical behaviour [44], [58]. Tree-based ensemble and neural-network approaches, including Random Forest [19], [59], XGBoost [50], [60], and ANN [17], [18], have been applied with varying degrees of success across different geological settings, yet their relative performance on GR-based prediction tasks has not been systematically benchmarked.

A particularly critical gap lies in the underutilization of gamma-ray logs as standalone predictors of rock mechanical properties. While Zhang et al. [58] and Horra et al. [38] have highlighted the broader potential of geophysical logs for mechanical property prediction, few studies have specifically investigated GR logs as a primary or sole predictor. This is surprising, given that GR measurements directly correlate with clay content and mineral composition – both of which are well-established controls on rock strength [41], [43]. The absence of systematic, GR-focused prediction studies represents a missed opportunity, particularly for mining operations where GR logs may be the most readily available continuous dataset.

This study addresses these gaps by re-evaluating the relationship between gamma-ray logs and fundamental rock mechanical parameters through a rigorous comparison of modelling approaches. Specifically, we aim to:

– determine whether the relationship between GR logs and mechanical properties is genuinely linear or whether non-linear patterns exist that simpler models might miss;

– identify which mechanical parameters UCS, cohesion, and friction angle show the strongest correlation with GR measurements;

– develop practical, field-deployable prediction equations that balance accuracy with simplicity for real-world applications.

Our contributions are threefold. First, we challenge the common assumption that complex ML models inherently outperform simpler approaches, demonstrating instead that linear models often yield more reliable and interpretable predictions a critical advantage in field practice where transparency is as important as accuracy. Second, we establish the first systematic empirical relationship between GR logs and multiple rock-mechanical parameters using a curated dataset of 54 depth-matched records, directly addressing the paucity of studies that employ GR as a primary predictor. Third, we derive regression equations that engineers can readily apply for rapid geomechanical assessment, reducing reliance on costly laboratory testing while maintaining scientific rigour. The novelty of this study lies in treating GR logs as a principal predictor of rock mechanical properties an approach rarely adopted in prior research, which has largely focused on sonic or density logs. By validating GR-based prediction through both linear and non-linear models, this work closes a critical research gap and introduces a transparent, field-deployable framework for geomechanical applications.

2. Methods

2.1. Data description

The dataset used in this research originates from the file logGRtoRM.xlsx, compiling well log measurements and corresponding laboratory test results from multiple boreholes, totaling 54 depth-specific records. The acquisition process began with field testing, specifically Gamma-Ray (GR) logging, during which the borehole location was determined based on geological objectives. The in-situ GR log measurement was acquired using calibrated wireline tools, in accordance with ASTM D6274 (Standard Guide for Conducting Borehole Geophysical Logging – Gamma) and API standards, ensuring data quality for lithological and clay content assessment.

Concurrently, core samples were retrieved from the logged intervals for subsequent laboratory testing. To preserve the rock original properties, the samples were handled with meticulous care: wrapped in plastic wrap and aluminum foil, secured within PVC pipes to maintain their natural moisture content and prevent defects during handling and transport.

These protected samples were then subjected to geomechanical testing in the laboratory following established standards. Rock mechanical properties were determined using methods such as ASTM D7012 (Standard Test Method for Compressive Strength and Elastic Moduli of Intact Rock Core Specimens). Specifically, the Uniaxial Compressive Strength (UCS) was measured in accordance with ISRM Suggested Methods, while cohesion and friction angle were derived from triaxial compression tests. Additional physical properties, including density, natural water content, and void ratio, were determined in accordance with ASTM D7012, ASTM D7263, and related geotechnical testing standards.

The compiled dataset included petrophysical, spatial, and geomechanical variables describing the studied formation. Gamma ray (GR) served as the main logging-derived predictor, whereas depth represented the position of measurement within the borehole. The remaining variables characterized the rock mass physical and mechanical properties, including natural water content, density, void ratio, friction angle, cohesion, and uniaxial compressive strength (UCS). The complete list of variables and their corresponding units is given in Table 1.

Table 1. Variables included in the dataset

Variable	Description	Unit
Gamma ray (GR)	Measurement of natural radioactivity; principal input feature	API
Depth	Measurement depth within the borehole	m
Natural water content	Natural moisture content of the formation	%
Density	Bulk density of rock specimens (gravimetric methods)	g/cm^3
Void ratio, *e*	Ratio of pore volume to solid volume in the rock matrix	–
Friction angle, φ	Internal angle of shear resistance from triaxial tests	°
Cohesion, *c*	Cohesive strength from Mohr-Coulomb failure envelope intercept	kPa
UCS	Maximum axial stress before failure in uniaxial compression	kPa

This structured acquisition ensures that both field-based log data and laboratory-based strength measurements are compatible, reliable, and aligned with internationally recognized standards, providing a robust foundation for the subsequent modeling and predictive analyses.

2.2. Correlation analysis and feature selection

Before model implementation, a comprehensive correlation analysis was performed to identify the most influential input features for predicting rock mechanical parameters. A correlation matrix was generated to quantify the linear relationships between all variables. This step is crucial for understanding the dataset inherent dependencies and guiding the feature selection process. The analysis revealed that while a set of variables, including Density, Natural Water Content, and Void Ratio, exhibited a strong correlation with the target parameters (UCS, cohesion, and friction angle), a specific focus was given to the gamma ray log due to its non-linear and complex relationship with rock strength. The subsequent models were built using this multi-feature dataset to leverage these identified relationships and improve predictive accuracy.

2.3. Data preprocessing

Effective data analysis requires thorough preprocessing, including data cleaning, handling missing values, feature selection, and feature scaling. The steps are as follows:

1. Data Cleaning. The raw dataset contains non-informative column names and missing entries. Column names are renamed for clarity, and missing values are imputed using median values to mitigate the influence of outliers.

2. Feature Scaling. To ensure uniformity and prevent features with larger ranges from dominating the learning process, all input features are standardized using the StandardScaler from the Scikit-learn library, setting each feature to have a mean of zero and a standard deviation of one.

2.4. Model implementation

After preprocessing, the dataset is split into training (80%) and test (20%) subsets. Four regression models are implemented to predict each mechanical parameter:

1. Linear Regression. A fundamental statistical model assuming linear relationships between variables, serving as a baseline for comparison.

2. Random Forest (RF) models have emerged as reliable tools in geomechanics for predicting rock mechanical properties, providing a data-driven alternative to traditional empirical formulas [35], [61]. As an ensemble of decision trees, an RF captures complex nonlinear relationships between input features (e.g., porosity, sonic velocity) and outputs like uniaxial compressive strength, often yielding high accuracy in rock property prediction. Studies report that RF bagging approach tends to reduce overfitting, and its feature importance metrics help engineers interpret which factors most influence the predicted strength. This combination of simplicity and robustness makes RF particularly suitable for rock mechanics prediction tasks. The main hyperparameters used in this study for Random Forest are summarized in Table 1. Additionally, the workflow architecture of Random Forest as applied in this study is presented in Figure 1.

Extreme Gradient Boosting (XGBoost) has proven to be a powerful machine learning method for rock mechanics prediction tasks, particularly for its high predictive performance.

Table 2. Random Forest hyperparameters used in this study

Hyperparameter	Value / setting
n_estimators	100
max_depth	None (nodes expanded until all leaves are pure)
min_samples_split	2
min_samples_leaf	1
max_features	auto (sqrt of total features)
bootstrap	True
random_state	42
criterion	Squared error (MSE)

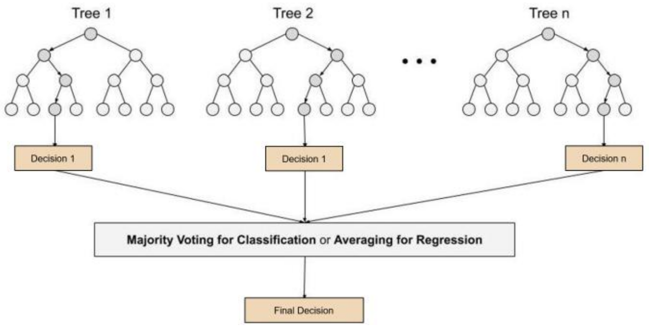

Figure 1. Conceptual workflow of Random Forest

This gradient boosting technique builds an ensemble of decision-tree learners sequentially, allowing it to capture subtle nonlinear trends in geomechanical datasets and to improve accuracy with each iteration. Recent studies in mining geomechanics have shown that XGBoost models can achieve excellent performance (often over 80% prediction accuracy) in estimating rock properties and classifying lithologies [53]. Its ability to handle limited, heterogeneous data and to perform strong regularization makes XGBoost well-suited for geotechnical applications, providing reliable predictions of key parameters such as rock UCS and Young's modulus [1]. The key hyperparameters applied in this study are presented in Table 3. Moreover, the sequential boosting architecture of XGBoost, as applied in this study, is shown in Figure 2.

Table 3. XGBoost hyperparameters used in this study

Hyperparameter	Value / setting
n_estimators	100
max_depth	6
learning_rate	0.1
subsample	0.8
colsample_bytree	0.8
reg_alpha (α)	0 (L1 regularization)
reg_lambda (λ)	1 (L2 regularization)
objective	reg:squarederror
random_state	42

Artificial Neural Networks (ANNs) have become a staple in geomechanics research for predicting rock mechanical properties due to their capacity to learn complex, nonlinear relationships from data. These models, inspired by biological neural systems, can integrate diverse inputs (e.g., density, porosity, seismic velocity) to estimate outputs like uniaxial compressive strength or elastic modulus with high accuracy [1]. In rock mechanics studies, even relatively simple feed-forward ANNs have demonstrated reliable performance; for instance, a sequential ANN model slightly outperformed other techniques in predicting the UCS of carbonate rocks from index tests [26].

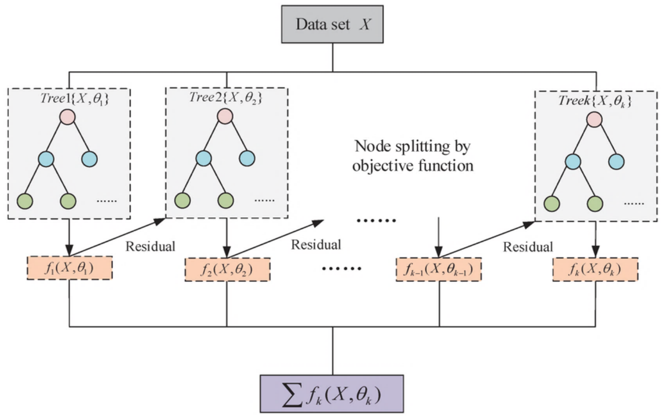

Figure 2. Conceptual workflow of XGBoost

By letting data-driven patterns inform their predictions, ANNs offer flexibility in modeling nonlinear data patterns but require careful configuration to avoid overfitting. The principal hyperparameters employed for the ANN model are summarized in Table 4. In addition, the multilayer perceptron architecture of the ANN model implemented in this study is depicted in Figure 3.

Table 4. ANN hyperparameters used in this study

Hyperparameter	Value / setting
Architecture	Sequential feed-forward (MLP)
Hidden layers	2 (64 neurons, 32 neurons)
Activation function	ReLU (hidden layers), linear (output)
Optimizer	Adam
Learning rate	0.001
Epochs	200
Batch size	16
Loss function	Mean squared error (MSE)
Early stopping	Patience = 20 epochs
Validation split	0.2 (from training set)

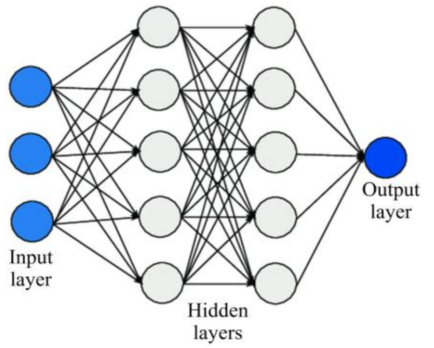

Figure 3. Conceptual workflow of ANN

2.5. Performance metrics

The predictive performance of each model is assessed using three standard regression metrics: the Coefficient of Determination (R^2), Mean Squared Error (MSE), and Root Mean Square Error (RMSE). The R^2 score quantifies the proportion of variance in the dependent variable explained by the model, with values closer to 1 denoting a superior fit. MSE and RMSE measure the average magnitude of the prediction error in squared units and the same units as the predicted variable, respectively. Model comparison is conducted based on these metrics obtained from the test dataset for each predicted parameter, facilitating an objective evaluation of model efficacy and robustness.

2.6. Model interpretability

To move beyond the "black box" nature of complex machine learning models, the study employed SHapley Additive exPlanations (SHAP) values to interpret the results of the best-performing models. This method was used to analyze the contribution and impact of each input feature on the model predictions, providing a deeper understanding of the relationships between the input logs and the predicted rock mechanical parameters.

3. Results and discussion

The dataset integrates geological and geomechanical variables that collectively capture lithological heterogeneity and rock strength behavior. Depth establishes the stratigraphic framework, while natural water content highlights the weakening effect of moisture on strength, consistent with experimental findings on water-induced reduction of brittle failure in rocks [3], [58]. Density and void ratio reflect compaction and pore structure, both of which are strongly linked to mechanical competence [56]. Variations in uniaxial compressive strength (UCS), cohesion, and friction angle illustrate the fundamental resistance of rocks to loading and shearing, parameters that remain central to geotechnical design and stability analyses [4]. Gamma-ray measurements, by capturing variability in mineralogy and clay content, further emphasize their utility as indirect indicators of mechanical behavior across sedimentary formations [61].

3.1. Descriptive statistics

The descriptive statistics provide a foundational understanding of the dataset structure. Table 5 below summarizes the central tendency, dispersion, and distributional shape of key variables, including Depth, Natural Water Content, Density, Void Ratio, Strength Type, UCS, Cohesion, Friction Angle, and gamma ray.

Table 5. Descriptive statistics of the dataset

Parameter	Depth, m	Natural water content, %	Density, gr/cm^3	Void ratio	Strength type	UCS, kPa	Cohesion, kPa	Friction angle, degree	Gamma ray, API
Mean	27.15	19.89	2.06	1.72	0.51	591.02	62.74	19.21	62.76
Standard error	2.38	0.61	0.02	0.02	0.01	36.31	1.42	0.55	1.76
Median	24.00	18.41	2.07	1.74	0.49	611.46	64.80	19.40	65.36
Standard deviation	17.34	4.45	0.12	0.12	0.11	264.37	10.32	3.98	12.84
Sample variance	300.55	19.77	0.01	0.01	0.01	69894.01	106.56	15.82	164.76
Kurtosis	-0.50	0.60	1.15	0.23	1.52	-0.43	-0.73	-0.64	-0.56
Skewness	0.55	0.91	-0.86	-0.64	1.06	-0.02	-0.50	-0.42	-0.50
Range	66.50	21.52	0.59	0.52	0.55	1130.30	36.36	15.71	51.78
Minimum	2.00	11.44	1.69	1.43	0.32	73.25	43.25	10.52	30.00
Maximum	68.50	32.96	2.28	1.95	0.87	1203.55	79.61	26.23	81.78
Count	53	53	53	53	53	53	53	53	53

Descriptive statistics reveal that the dataset spans a wide depth range (2.00-68.50 m) with a mean of 27.15 m, indicating substantial vertical variability in the sampled formation. Natural water content averages 19.89% (SD = 4.45%), while density (mean = 2.06 g/cm^3) and void ratio (mean = 1.72) show relatively low dispersion, suggesting consistent compaction. The uniaxial compressive strength (UCS) exhibits high variability (mean = 591.02 kPa, SD = 264.37 kPa), reflecting the heterogeneous nature of the sedimentary rock. Cohesion and friction angle display moderate values (62.74 kPa and 19.21°, respectively), typical of fine-grained soils. Skewness and kurtosis values indicate non-normal distributions for several parameters, particularly UCS and strength type, suggesting asymmetry and potential outliers, which may influence geotechnical interpretation. Gamma ray readings (mean = 62.76 API) further support lithological heterogeneity across the profile. Overall, the data highlight significant spatial variation in mechanical and physical properties, underscoring the need for detailed site-specific characterization.

The results section summarizes two complementary analyses. First, Pearson's correlation coefficients were computed to explore linear relationships among all variables. Second, three machine-learning models, Random Forest (RF), Extreme Gradient Boosting (XGBoost) and an Artificial Neural Network (ANN) were tuned to predict each mechanical property from the available log data. The discussion interprets both the correlation patterns and the predictive performance of the models.

3.2. Visualization and discussion of rock mechanics properties

The column charts, which profile the five critical subsurface properties across depth, immediately reveal coherent, often parallel trends that serve as the geological "fingerprints" of the area. These consistent patterns strongly suggest potential correlations among the parameters, providing a powerful initial insight.

Crucially, as shown in Figure 4, changes in lithology-dependent parameters (gamma ray and density) frequently coincide with shifts in rock mechanical properties (Friction Angle, Cohesion, and UCS). For example, in borehole E, a notable increase in both gamma ray and density at a shallow depth of approximately 5.7 m is immediately followed by a corresponding increase in the mechanical variables (friction angle, cohesion, and UCS). This overall, synchronized movement, in which the lithological shift signals a change in the mechanical strength profile, is a recurring observation.

More broadly, a similar overall trend is observable between the gamma ray profile and the rock mechanics properties across the full depth range. Corresponding increases often mirror the rising and falling values in gamma ray, while decreases in friction angle, cohesion, and UCS are observed. This consistent, depth-wise synchronization, such as the general decrease followed by an increase observed between 20.5 and 36.3 m, implies that the mechanical competence of the rock is fundamentally linked to its geological composition, likely driven by variations in clay content or matrix density.

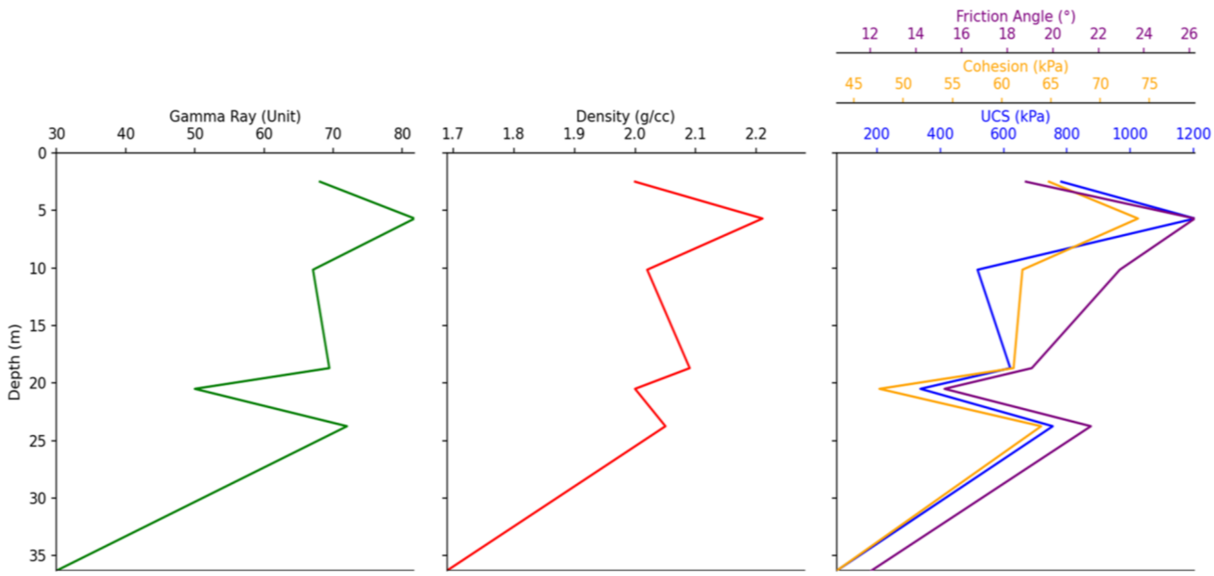

Figure 4. Composite log showing gamma ray, density, UCS, cohesion, and friction angle for borehole E

This recognition of powerful visual correlations is the essential first step, guiding our subsequent detailed quantitative analyses and paving the way for more accurate geotechnical and geological evaluations across all 54 datasets.

3.3. Correlation between variables

Table 6 reports on the Pearson correlation coefficients between each pair of variables. Strong positive or negative correlations are highlighted. As expected, uniaxial compressive strength (UCS), cohesion and friction angle are strongly intercorrelated ($r \approx 0.78$-0.83) because they all reflect rock strength. Gamma-ray shows high positive correlation with UCS (0.852), cohesion (0.799) and friction angle (0.785), indicating that higher radioactivity (often associated with shallower formations) corresponds to stronger rock in this dataset. Natural water content, by contrast, has strong negative correlations with all three mechanical parameters ($r \approx -0.65$ to -0.76), reflecting the weakening effect of moisture. Density and void ratio correlate positively with UCS and cohesion, whereas the categorical strength type variable correlates negatively with them. Depth has only weak positive correlations with the mechanical properties, suggesting that burial depth alone does not strongly control strength in these samples.

Table 6. Pearson correlation matrix

Variables	Depth	Natural water content	Density	Void ratio	Strength type	UCS	Cohesion	Friction angle	Gamma ray
Depth	1.000	-0.253	0.206	0.265	-0.248	0.305	0.328	0.244	0.319
Natural water content	-0.253	1.000	-0.695	-0.827	0.705	-0.757	-0.653	-0.710	-0.813
Density	0.206	-0.695	1.000	0.707	-0.823	0.773	0.620	0.637	0.743
Void ratio	0.265	-0.827	0.707	1.000	-0.772	0.791	0.660	0.691	0.725
Strength type	-0.248	0.705	-0.823	-0.772	1.000	-0.766	-0.710	-0.676	-0.733
UCS	0.305	-0.757	0.773	0.791	-0.766	1.000	0.783	0.797	0.852
Cohesion	0.328	-0.653	0.620	0.660	-0.710	0.783	1.000	0.819	0.799
Friction angle	0.244	-0.710	0.637	0.691	-0.676	0.797	0.819	1.000	0.785
Gamma ray	0.319	-0.813	0.743	0.725	-0.733	0.852	0.799	0.785	1.000

The strong positive correlations (bold) indicate that gamma-ray, density and void ratio tend to increase together with UCS and cohesion, while natural water content and strength type decrease with them. These patterns imply that combining multiple log features could help machine-learning models capture the complex controls on rock strength.

3.4. Machine-learning model performance

To quantify predictive capability, Random Forest, XGBoost and ANN models were tuned for each target variable. All models used the full set of predictors (gamma-ray, depth, natural water content, density, void ratio and strength type). Hyperparameters were optimized via grid search; the optimal settings and resulting test-set metrics are listed in Table 2 along with concise interpretations.

To quantify predictive capability, Random Forest, XGBoost, and ANN models were tuned for each target variable. All models used the full set of predictors (gamma ray, depth, natural water content, density, void ratio, and strength type). Hyperparameters were optimized via grid search; the optimal settings and resulting test-set metrics are listed in Table 7.

3.4.1. Cohesion

The correlation matrix indicates that cohesion is strongly associated with gamma-ray, UCS and friction angle ($r \approx 0.80$-0.82), as well as density and void ratio ($r \approx 0.62$-0.66). XGBoost leverages these relationships by sequentially reducing residual errors, leading to an R^2 of 0.6764 and the lowest RMSE (5.09 kPa). Random Forest attains $R^2 \approx 0.51$, reflecting its ability to model moderate nonlinearities through bagging but not fully capturing higher-order interactions.

Table 7. Performance of machine learning models for each target variable

Target	Model	R^2	RMSE	Interpretation
Cohesion	Random Forest	0.5094	6.27 kPa	The ensemble of unpruned decision trees captures moderate nonlinear patterns and explains over half of the variance.
	XGBoost	0.6764	5.09 kPa	Gradient boosting exploits complex feature interactions and provides the best overall performance for cohesion
	ANN	0.3741	7.08 kPa	The ANN struggles with the limited dataset and tends to overfit, leading to lower performance than that of tree-based models.
Friction angle	Random Forest	0.4955	2.35°	Bagging produces the most accurate estimates for the friction angle and captures roughly half of the variance.
	XGBoost	0.3550	2.66°	Boosting reduces bias but may overfit on the small dataset, resulting in lower performance than RF.
	ANN	0.2727	2.82°	The neural network fails to learn robust patterns from the available tabular log data.
UCS	Random Forest	0.4200	148.35 MPa	RF explains about 42% of the UCS variance, although residual errors remain substantial due to the complex nature of rock behavior.
	XGBoost	0.4427	145.42 MPa	XGBoost slightly outperforms RF by capturing nonlinear relationships among the predictors.
	ANN	0.3838	152.91 MPa	The ANN shows a marginal improvement over the baseline but still lags behind the tree-based models.

The ANN, despite its flexibility, explains less than 38% of the variance in cohesion; this likely results from limited training samples and the absence of highly nonlinear patterns that would justify a neural network.

3.4.2. Friction angle

Friction angle correlates strongly with cohesion and UCS ($r \approx 0.82$) and moderately with gamma-ray, density and void ratio ($r \approx 0.69$-0.79). In the tuned experiments, Random Forest achieved the highest R^2 (0.4955) and lowest RMSE (2.35°), suggesting that the hierarchical splits in the ensemble of trees effectively capture the dependence of friction angle on multiple physical logs. XGBoost performed slightly worse ($R^2 \approx 0.3550$), perhaps because boosting introduces additional variance on the small dataset. The ANN provided the least accurate predictions ($R^2 \approx 0.2727$), indicating that the friction angle may depend on subtle lithological features not represented in the available variables.

3.4.3. Uniaxial compressive strength (UCS)

UCS exhibits very strong positive correlations with gamma-ray, density and void ratio ($r \approx 0.77$-0.85) and negative correlations with natural water content and strength type. Both Random Forest and XGBoost capture around 42-44 % of the variance, with XGBoost slightly outperforming RF. The high RMSE values (~145-148 MPa) indicate that substantial unexplained variability remains; this is not surprising given that UCS depends on micro-scale fabric, mineralogy and fracture networks that are not directly sensed by basic log measurements. The ANN again underperforms due to the small dataset and the lack of more complex patterns to exploit.

3.5. Summary of model selection

Table 8 summarizes the best model for each target variable based on R^2 and RMSE, along with interpretative remarks. XGBoost is recommended for cohesion and UCS, whereas Random Forest is preferable for friction angle. Exponential regression failed to converge and thus is not considered. The ANN did not surpass tree-based models in any case.

3.6. Discussion

The descriptive statistics reveal marked heterogeneity in the dataset, particularly for UCS (mean = 591.02 kPa, SD = 264.37 kPa), which reflects the variable strength of the sampled sedimentary formations. Natural water content also shows considerable variability, averaging 19.89% and exhibiting a strong negative correlation with strength parameters, consistent with the known weakening effect of pore fluids. By contrast, density and void ratio exhibit lower dispersion, suggesting more uniform compaction across the profile. Gamma-ray readings (mean = 62.76 API) highlight lithological variability, which is strongly associated with UCS, cohesion, and friction angle, further supporting this association.

Correlation analysis confirms that gamma-ray, density, and void ratio are positively associated with mechanical strength, whereas natural water content is negatively associated. The high correlation of gamma-ray with UCS ($r = 0.852$), cohesion ($r = 0.799$), and friction angle ($r = 0.785$) suggests that radioactivity, likely reflecting clay content and lithological changes, provides a useful proxy for mechanical behavior. Nevertheless, the predictive modeling demonstrates that gamma-ray alone is insufficient to capture the full complexity of rock strength. Models incorporating multiple log variables yield superior performance, underscoring the multivariate nature of geomechanical controls.

Among the machine-learning approaches, XGBoost achieved the best performance for cohesion ($R^2 = 0.6764$, RMSE = 5.09 kPa) and UCS ($R^2 = 0.4427$, RMSE = 145.42 MPa), reflecting its ability to model nonlinear interactions between gamma-ray, density, void ratio, and moisture (Fig. 5). Random Forest provided the most accurate estimates of friction angle ($R^2 = 0.4955$, RMSE = 2.35°), benefiting from its robustness against overfitting and its capability to capture hierarchical dependencies. In contrast, the ANN consistently underperformed, with R^2 values below 0.40 and, in some cases, negative, indica-ting overfitting and limited suitability for small tabular datasets. These outcomes align with previous findings that tree-based ensembles outperform neural networks when sample sizes are small, and the feature space is relatively constrained.

Several critical limitations are worth discussing. First, the dataset comprises only 54 depth-matched records, constraining the training of data-intensive models such as ANNs and limits the statistical power of cross-validation. The 80/20 train-test split, while standard, yields a test set of approximately 11 samples, a size at which R^2 estimates may be unstable.

Table 8. Summary of optimal models and recommendations

Target	Best model	R^2	RMSE	Comment
Cohesion	XGBoost	0.6764	5.09 kPa	The model captures complex interactions among gamma-ray, density, void ratio, and moisture and is the most accurate for cohesion.
Friction angle	Random Forest	0.4955	2.35°	The ensemble of trees best models moderate nonlinear relationships and remains robust to overfitting on small datasets.
UCS	XGBoost	0.4427	145.42 MPa	The model slightly outperforms RF; however, the high error indicates the need for additional logs or micro-scale data.

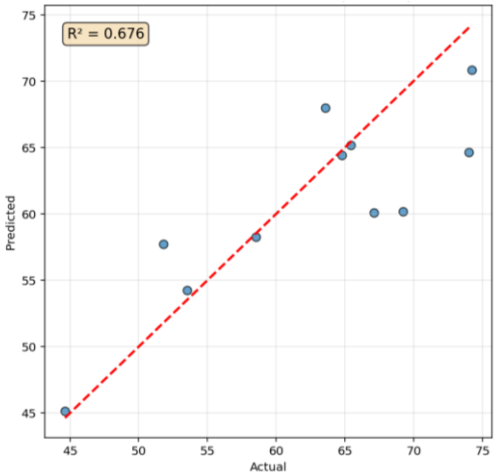

Figure 5** **Comparison of predicted and actual cohesion values using the XGBoost model (R^2 = 0.676)

Future studies should adopt *k*-fold cross-validation or leave-one-out cross-validation to provide more robust performance estimates. Second, the explained variance remains insignificant across all targets: less than half of the variability in UCS ($R^2 = 0.4427$) and friction angle ($R^2 = 0.4955$) is captured by the best models. This indicates that key determinants of rock strength, such as mineral composition, cementation degree, grain-size distribution, and microfracture density, are not captured by standard logging measurements. The inclusion of additional well logs (e.g., sonic, resistivity, neutron) and petrographic or mineralogical data would likely improve predictive capability.

Third, the generalizability of the derived regression equations is inherently constrained by the dataset lithological and geographical scope. The models were trained on sedimentary formations from a specific study area, and their transferability to other lithologies or geological settings has not been validated. Caution is therefore advised when applying these equations beyond the calibration domain. Fourth, while the SHAP analysis provides valuable insight into feature contributions, its interpretation should be contextualized by the limited dataset size; SHAP values derived from small samples may not fully represent the underlying feature-target relationships at the population level.

Despite these limitations, the study makes a meaningful contribution by demonstrating that GR logs can serve as a practical, cost-effective proxy for rock mechanical properties. The finding that linear regression performs comparably to, or even better than, more complex ML models is particularly significant: it suggests that the GR mechanical property relationship is predominantly linear at the resolution of standard wireline data, and that simple, transparent models may be preferable for field deployment, where interpretability and reproducibility are essential.

Future research should focus on expanding the dataset across diverse lithologies and geological settings to improve generalizability, integrating additional log types such as sonic, density, neutron, and resistivity logs for multivariate prediction, incorporating mineralogical and petrographic data to capture micro-scale controls on rock strength, applying advanced cross-validation and ensemble strategies to maximize model robustness, and exploring transfer learning and domain adaptation techniques to extend model applicability across geological domains.

4. Conclusions

This study demonstrated the feasibility of using gamma ray (GR) logs to estimate key rock mechanical parameters, cohesion, internal friction angle, and UCS by comparing regression and machine learning approaches on a depth-matched dataset of 54 records from sedimentary formations. The principal findings and conclusions are as follows. First, the GR-mechanical property relationship is largely linear at the resolution of wireline data, with linear regression performing comparably to, or even better than, more complex ML models. This finding challenges the common assumption that complex models inherently yield superior performance, and underscores the value of transparent, interpretable approaches in field practice.

Second, among machine learning methods, XGBoost provided the best predictions for cohesion ($R^2 = 0.6764$) and UCS ($R^2 = 0.4427$), while Random Forest was optimal for friction angle ($R^2 = 0.4955$). ANN consistently underperformed, confirming its unsuitability for small tabular datasets without careful regularization and architecture optimization.

Third, GR exhibited the strongest associations with UCS ($r = 0.852$) and cohesion ($r = 0.799$), reinforcing its role as a practical proxy for rock strength. However, multivariate inputs combining GR with density, water content, and void ratio remain essential for stable predictions.

Fourth, the study highlights that simple linear models can deliver transparent and field-ready regression equations, offering a cost-effective alternative to extensive laboratory testing. Nevertheless, unresolved influences from mineralogy, cementation, grain size, and microfractures restrict predictive accuracy, with less than half of the variance explained for UCS and friction angle.

Future work should focus on enlarging datasets across different lithologies, integrating mineralogical data, applying rigorous cross-validation strategies, and exploring transfer learning techniques to enhance generalizability. Collectively, these steps will strengthen GR-based prediction frameworks, balancing accuracy and transparency for geomechanical applications in mining and engineering practice.

Author contributions

Conceptualization: ABS, SS, BD; Data curation: ABS, SS, RS; Formal ana-lysis: SS, MAF; Investigation: SS, MAF; Methodology: ABS, SS, RS, BD; Project administration:

ABS, RS, BR; Resources: ABS, SS; Software: MAF; Supervision: SS, BR; Validation: SS, MAF; Visualization: MAF; Writing – original draft: ABS, MAF, RS; Writing – review & editing: ABS, SS, RS, BR. All authors have read and agreed to the published version of the manuscript.

Funding

This research received no external funding.

Acknowledgements

The author would like to thank the Institut Teknologi Sains Bandung (ITSB), Institut Teknologi Nasional Yogyakarta (ITNY), and Universitas Pembangunan Nasional Veteran Yogyakarta (UPNVY) for their support of this research.

Conflicts of interests

The authors declare no conflict of interest.

Data availability statement

The original contributions presented in the study are included in the article, further inquiries can be directed to the corresponding author.

References

[1] Shahani, N.M., Ullah, B., Shah, K.S., Hassan, F.U., Ali, R., Elkotb, M.A., Ghoneim, M.E., & Tag-Eldin, E.M. (2022). Predicting angle of internal friction and cohesion of rocks based on machine learning algorithms. *Mathematics*, *10*(20), 3875. https://doi.org/10.3390/math10203875

[2] Deniz, B.E., & Topal, T. (2023). Prediction of uniaxial compressive strength of the Kızılkaya ignimbrite with variable properties using MRA and ANN, Cappadocia (Turkey). *Discover Environment*, *1*(1), 12. https://doi.org/10.1007/s44274-023-00013-1

[3] Bieniawski, Z.T. (1974). Estimating the strength of rock materials. *Journal of the South African Institute of Mining and Metallurgy*, *74*(9), 312-320.

[4] Aladejare, A.E. (2020). Evaluation of empirical estimation of uniaxial compressive strength of rock using measurements from index and physical tests. *Journal of Rock Mechanics and Geotechnical Engineering*, *12*, 256-268. https://doi.org/10.1016/j.jrmge.2019.08.001

[5] Aladejare, A.E., Alofe, E.D., Onifade, M., Lawal, A.I., Ozoji, T.M., & Zhang, Z.-X. (2021). Empirical estimation of uniaxial compressive strength of rock: Database of simple, multiple, and artificial intelligence-based regressions. *Geotechnical and Geological Engineering*, *39*(6), 4427-4455. https://doi.org/10.1007/s10706-021-01772-5

[6] Madhubabu, T.V., Murthy, Y.R., & Hari Prasad, K. (2016). Prediction of compressive strength and elastic modulus of carbonate rocks. *Measurement*, *93*, 519-526. https://doi.org/10.1016/j.measurement.2016.03.050

[7] Rahman, T., & Sarkar, K. (2021). Lithological control on the estimation of uniaxial compressive strength by the P-wave velocity using supervised and unsupervised learning. *Rock Mechanics and Rock Engineering*, *54*(6), 3175-3191. https://doi.org/10.1007/s00603-021-02445-8

[8] Shakoor, A., & Bonelli, R.E. (1991). Relationship between petrographic characteristics, engineering index properties and mechanical properties of selected sandstones. *Bulletin of the Association of Engineering Geologists*, *28*(1), 55-71. https://doi.org/10.2113/gseegeosci.xxviii.1.55

[9] Teymen, A., & Mengüç, A. (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.

[10] Jeng, F.S., Lin, M.L., Huang, T.H., & Weng, M.C. (2004). Influence of petrographic parameters on geotechnical behavior of weak sandstone. *Engineering Geology*, 73(1-2), 71-91. https://doi.org/10.1016/j.enggeo.2003.12.001

[11] Duda, M. (2013). The weakening effect of water on the brittle failure strength of rocks. *Geophysical Journal International*, *192*(3), 1091-1108. https://doi.org/10.1093/gji/ggs090

[12] Schumacher, R., & Mues-Schumacher, U. (1996). The Kızılkaya ignimbrite – An unusual low-aspect-ratio ignimbrite from Cappadocia, central Turkey. *Journal of Volcanology and Geothermal Research*, *70*(-2), 107-121. https://doi.org/10.1016/0377-0273(95)00046-1

[13] Korkanç, M. (2016). Estimation of engineering properties of selected tuffs by petrographic and rock indices. *Journal of African Earth Sciences*, *120*, 12-23. https://doi.org/10.1016/j.jafrearsci.2016.05.008

[14] Zhang, X., Altalbawy, F.M.A., Gasmalla, T.A.S., Al-Khafaji, A.H.D., Iraji, A., Syah, R.B.Y., & Nehdi, M.L. (2023). Performance of statistical and intelligent methods in estimating rock compressive strength. *Sustainability*, *15*(7), 5642. https://doi.org/10.3390/su15075642

[15] Liu, Z., Li, D., Liu, Y., Yang, B., & Zhang, Z.-X. (2023). Prediction of uniaxial compressive strength of rock based on lithology using stacking models. *Rock Mechanics Bulletin*, *2*(4), 100081. https://doi.org/10.1016/j.rockmb.2023.100081

[16] Asteris, P.G., Karoglo, M., Skentou, M., Vasconcelos, G.D., Mingming, H., Bakoulas, G., & Tsakiridis, P. (2024). Predicting uniaxial compressive strength of rocks using ANN models: Incorporating porosity, compressional wave velocity, and Schmidt hammer data. *Ultrasonics*, *130*, 107347. https://doi.org/10.1016/j.ultras.2024.107347

[17] Ferentinou, M., & Fakir, M. (2017). An ANN approach for the prediction of uniaxial compressive strength of some sedimentary and igneous rocks in Eastern KwaZulu-Natal. *Procedia Engineering*, *191*, 1117-1125. https://doi.org/10.1016/j.proeng.2017.05.286

[18] Wei, X., Shahani, N.M., & Zheng, X. (2023). Predictive modeling of the uniaxial compressive strength of rocks using an artificial neural network approach. *Mathematics*, *11*(7), 1650. https://doi.org/10.3390/math11071650

[19] Wang, M., Wan, W., & Zhao, Y. (2020). Prediction of the uniaxial compressive strength of rocks from simple index tests using a random forest predictive model. *Comptes Rendus Mécanique*, *348*(1), 3-32. https://doi.org/10.5802/crmeca.3

[20] Xu, W., Wang, M., Liu, W., Liu, H., Li, H., & Wang, Q. (2024). Comparative study on convolutional neural network and regression analysis to evaluate uniaxial compressive strength of sandy dolomite. *Scientific Reports*, *14*, 9880. https://doi.org/10.1038/s41598-024-60085-8

[21] Zheng, X., Batool, A., Kumar, S., & Shahani, N.M. (2026). Grey Wolf Optimization-optimized ensemble models for predicting the uniaxial compressive strength of rocks. *Applied Sciences*, *16*(2), 1130. https://doi.org/10.3390/app16021130

[22] Cao, J., Gao, J., Nikfashan Rad, H., Mohammed, A. S., Hasanpanah, M., Zhou, J., & Li, Y. (2022). A novel systematic and evolved approach based on XGBoost-firefly algorithm to predict Young's modulus and unconfined compressive strength of rock. *Engineering with Computers*, *38*(4), 1657-1673. https://doi.org/10.1007/s00366-020-01241-2

[23] Wang, M., Zhao, G., Liang, W., & Wang, N. (2023). A comparative study on the development of hybrid SSA-RF and PSO-RF models for predicting the uniaxial compressive strength of rocks. *Case Studies in Construction Materials*, *18*, e02191. https://doi.org/10.1016/j.cscm.2023.e02191

[24] Li, D., Liu, Z., Armaghani, D.J., Xiao, P., & Zhou, J. (2022). Novel ensemble intelligence methodologies for rockburst assessment in complex and variable environments. *Scientific Reports*, *12*(1), 1844. https://doi.org/10.1038/s41598-022-05844-3

[25] Liu, Z., Jahed Armaghani, D., Fakharian, P., Li, D., & Vladimirovich Ulrich, D. (2022). Rock strength estimation using several tree-based ML techniques. *Computer Modeling in Engineering & Sciences*, *133*(3), 799-824. https://doi.org/10.32604/cmes.2022.021165

[26] Hassan, M.Y., & Arman, H. (2022). Several machine learning techniques comparison for the prediction of the uniaxial compressive strength of carbonate rocks. *Scientific Reports*, *12*, 20969. https://doi.org/10.1038/s41598-022-25633-0

[27] Abdelhedi, M., Jabbar, R., Ben Said, A., Fetais, N., & Abbes, C. (2023). Machine learning for prediction of the uniaxial compressive strength within carbonate rocks. *Earth Science Informatics*, *16*(2), 1633-1649. https://doi.org/10.1007/s12145-023-00977-z

[28] Mohamad, E.T., Jahed Armaghani, D., Momeni, E., & Alavi Nezhad Khalil Abad, S.V. (2015). Prediction of the unconfined compressive strength of soft rocks: A PSO-based ANN approach. *Bulletin of Engineering Geology and the Environment*, *74*(3), 745-757. https://doi.org/10.1007/s10064-014-0638-0

[29] Wei, M., Meng, W., Dai, F., & Wu, W. (2022). Application of machine learning in predicting the rate-dependent compressive strength of rocks. *Journal of Rock Mechanics and Geotechnical Engineering*, *14*(5), 1356-1365. https://doi.org/10.1016/j.jrmge.2022.01.002

[30] Jang, J.S. (1993). ANFIS: Adaptive-network-based fuzzy inference system. *IEEE Transactions on Systems, Man, and Cybernetics*, *23*(3), 665-685. https://doi.org/10.1109/21.256541

[31] Jang, J.S., & Sun, C.T. (1995). Neuro-fuzzy modeling and control. *Proceedings of the IEEE*, *83*(3), 378-406. https://doi.org/10.1109/5.364486

[32] Al-Hamed, S., Wahby, M., Al-Sulaiman, M., & Aboukarima, A. (2014). Prediction of soil fractions (sand, silt and clay) in the surface layer based on natural radionuclide concentration in the soil using an adaptive neuro-fuzzy inference system. *Open Journal of Soil Science*, *4*(6), 215-224. https://doi.org/10.4236/ojss.2014.47024

[33] Cabalar, A.F., Cevik, A., Gokceoglu, C., & Baykal, G. (2010). Neuro-fuzzy based constitutive modeling of undrained response of Leighton

Buzzard Sand mixtures. *Expert Systems with Applications*, *37*(1), 842-851. https://doi.org/10.1016/j.eswa.2009.05.085
[34] Cao, W. (2025). A comparative study of hybrid adaptive neuro-fuzzy inference systems to predict the unconfined compressive strength of rocks. *Journal of Engineering and Applied Science*, *72*(1). https://doi.org/10.1186/s44147-024-00574-9
[35] Xie, Y., Li, X., & Min, Z. (2025). Comparison of machine learning models for rock UCS prediction using measurement while drilling data. *Scientific Reports*, *15*(1), 3456. https://doi.org/10.1038/s41598-025-93111-4
[36] Ali, M., & Lai, S.H. (2023). Artificial intelligent techniques for prediction of rock strength and deformation properties – A review. *Structures*, *55*, 1542-1555. https://doi.org/10.1016/j.istruc.2023.06.131
[37] Lawal, A. I., & Kwon, S. (2021). Application of artificial intelligence to rock mechanics: An overview. *Journal of Rock Mechanics and Geotechnical Engineering*, *13*(1), 248-266. https://doi.org/10.1016/j.jrmge.2020.05.010
[38] Horra, H., Hadjadj, A., Saad, E.A., & Brahim, K.M. (2025). 3D rock strength prediction by an innovative approach that integrates geostatistics with machine deep learning models. *Journal of Petroleum Exploration and Production Technology*, *15*(7), 112. https://doi.org/10.1007/s13202-025-02017-4
[39] Liang, J., Zhang, Y., Chen, H., & Liu, X. (2024). Analysis of rock mechanical characteristics and establishment of an evaluation system for coal measures in the Longtan Formation, Guizhou Province. *ACS Omega*, *9*(4), 3641-3654. https://doi.org/10.1021/acsomega.4c01643
[40] Mahmoodzadeh, A., Mohammadi, M., Ibrahim, H.H., Abdulhamid, S.N., Salim, S.G., Ali, H.F.H., & Majeed, M.K. (2021). Artificial intelligence forecasting models of uniaxial compressive strength. *Transportation Geotechnics*, *27*, 100499. https://doi.org/10.1016/j.trgeo.2020.100499
[41] Kumar, T., Seelam, N.K., & Rao, G.S. (2022). Lithology prediction from well log data using machine learning techniques: A case study from Talcher coalfield, Eastern India. *Journal of Applied Geophysics*, *199*, 104605. https://doi.org/10.1016/j.jappgeo.2022.104605
[42] Muzamhindo, H., & Ferentinou, M. (2023). Generic compressive strength prediction model applicable to multiple lithologies based on a broad global database. *Probabilistic Engineering Mechanics*, *71*, 103400. https://doi.org/10.1016/j.probengmech.2022.103400
[43] Sujatono, S. (2022). Determination of cohesion and friction angle on sedimentary rock based on geophysical log. *Geomechanics and Geophysics for Geo-Energy and Geo-Resources*, *8*(1), 33. https://doi.org/10.1007/s40948-022-00343-z
[44] Lai, J., Su, Y., Xiao, L., Zhao, F., Bai, T., Li, Y., Li, H., Huang, Y., Wang, G., & Qin, Z. (2024). Application of geophysical well logs in solving geologic issues: Past, present and future prospect. *Geoscience Frontiers*, *15*(3), 101779. https://doi.org/10.1016/j.gsf.2024.101779
[45] McNally, G.H. (1987). Estimation of coal measures rock strength using sonic and neutron logs. *Geoexploration*, *24*(4-5), 381-395. https://doi.org/10.1016/0016-7142(87)90008-1
[46] McNally, G.H. (1990). The prediction of geotechnical rock properties from sonic and neutron logs. *Exploration Geophysics*, *21*(1-2), 65-70. https://doi.org/10.1071/EG990065
[47] Saedi, B., Mohammadi, S., Davoud, S., & Hossein, S. (2018). Prediction of uniaxial compressive strength and elastic modulus of migmatites using various modeling techniques. *Arabian Journal of Geosciences*, *11*(18), 456. https://doi.org/10.1007/s12517-018-3912-9
[48] Miah, M.I., Ahmed, S., Zendehboudi, S., & Butt, S. (2020). Machine learning approach to model rock strength: Prediction and variable selection with aid of log data. *Rock Mechanics and Rock Engineering*, *53*, 4691-4715. https://doi.org/10.1007/s00603-020-02184-2
[49] Zhao, R., Shi, S., Li, S., Guo, W., Zhang, T., Li, X., & Lu, J. (2023). Deep learning for intelligent prediction of rock strength by adopting measurement while drilling data. *International Journal of Geomechanics*, *23*(4), 04023028. https://doi.org/10.1061/IJGNAI.GMENG-8133
[50] Xie, H., Lin, P., Kang, J., Zhai, C., & Du, Y. (2024). Prediction method of rock uniaxial compressive strength based on feature optimization and SSA-XGBoost. *Sustainability*, *16*(19), 8460. https://doi.org/10.3390/su16198460
[51] Li, D., Li, X., Liu, L., He, W., Li, Y., & Shi, S. (2025). Prediction on rock strength by mineral composition from machine learning of ECS logs. *Energy Geoscience*, *6*(2), 100386. https://doi.org/10.1016/j.engeos.2025.100386
[52] Goldstein, D., Aldrich, C., Shao, Q., & O'Connor, L. (2025). A machine learning classification approach to geotechnical characterization using measure-while-drilling data. *Geosciences*, *15*(3), 93. https://doi.org/10.3390/geosciences15030093
[53] Balaguera, A., Torné, M., Carbonell, R., Martí, A., Vergés, J., Jurado, M.J., & Rodríguez, S. (2025). Machine learning in subsurface physical properties and lithofacies prediction in a mining context. *Scientific Reports*, *15*, 26495. https://doi.org/10.1038/s41598-025-11953-4
[54] Sun, Q., Zhao, Y., & Ma, J. (2021). Predicting strength of coal – grout composites using ensemble machine learning methods. *International Journal of Mining Science and Technology*, *31*(2), 277-284. https://doi.org/10.1016/j.ijmst.2020.09.002
[55] Zhao, J., Li, D., Jiang, J., & Luo, P. (2024). Uniaxial compressive strength prediction for rock material in deep mine using boosting-based machine learning methods and optimization algorithms. *Computer Modeling in Engineering & Sciences*, *140*(1), 275-304. https://doi.org/10.32604/cmes.2024.046960
[56] Sujatono, S., & Wijaya, A. E. (2022). The influence of quartz content on modulus of elasticity and Poisson's ratio in quartz sandstone. *Bulletin of Engineering Geology and the Environment*, *81*(7), 292. https://doi.org/10.1007/s10064-022-02798-6
[57] Sujatono, S. (2022). Empirical analysis of density value based on laboratory test and geophysical log on sedimentary rock. *Geomate Journal*, *23*(98), 212-219. https://doi.org/10.21660/2022.98.3624
[58] Zhang, J., Liu, G., Wei, Z., Li, S., Zayier, Y., & Cheng, Y. (2025). Machine learning-based prediction of well logs guided by rock physics and its interpretation. *Sensors*, *25*(3), 836. https://doi.org/10.3390/s25030836
[59] Akosah, S., Gratchev, I., & Gidigasu, S.S.R. (2025). A systematic literature review on the application of artificial intelligence techniques for rock strength estimation. *Neural Computing and Applications*, 37, 20721-20753. https://doi.org/10.1007/s00521-025-11517-7
[60] Hussain, J., Fu, X., Chen, J., Ali, N., Tobal, S.M., Swamy, V.S., Kumar, M.B., & Karrar, R.C. (2025). Estimation of rock strength parameters from petrological contents using tree-based machine learning techniques. *AI in Civil Engineering*, *1*(1), 100001. https://doi.org/10.1007/s43503-024-00047-1
[61] Alomari, E., Ng, K., & Khatri, L. (2025). Experimental investigation and prediction of mechanical properties of carbonate rocks under uniaxial and triaxial compressions. *Materials*, *18*(6), 1211. https://doi.org/10.3390/ma18061211

Прогнозування механічних параметрів гірських порід за даними гамма-каротажу із використанням регресійних підходів і методів машинного навчання

А.Б. Сантосо, С. Суджатоно, М.А. Фархан, Р. Сяпутра, Б. Рахмад, Б. Двінагара

Мета. Дослідження доцільності використання даних гамма-каротажу (Gamma Ray, GR), отриманих кабельним каротажем, як основного предиктора для визначення ключових механічних параметрів гірських порід, зокрема кута внутрішнього тертя (φ), зчеплення (c) та межі міцності на одноосьовий стиск (UCS), як швидкої та економічно ефективної альтернативи традиційним лабораторним випробуванням.

Методика. Сформовано набір даних із 54 записів, узгоджених за глибиною, який містив значення гамма-каротажу та механічні властивості, визначені лабораторним шляхом. Дані отримано з кількох свердловин у межах осадових формацій. Прогностичні моделі побудовано з використанням лінійної регресії, Random Forest (RF), Extreme Gradient Boosting (XGBoost) та штучних нейронних мереж (ANN). Попередня обробка даних включала заповнення пропущених значень за медіаною, масштабування ознак та поділ вибірки на навчальну і тестову у співвідношенні 80/20. Ефективність моделей оцінювали за коефіцієнтом детермінації (R^2), середньоквадратичною похибкою (MSE) та коренем середньоквадратичної похибки (RMSE). Для покращення інтерпретованості моделей використано значення адитивних пояснень Шеплі (SHapley Additive exPlanations, SHAP).

Результати. Кореляційний аналіз виявив сильні додатні зв'язки між значеннями GR та UCS ($r = 0.852$), зчепленням ($r = 0.799$) і кутом внутрішнього тертя ($r = 0.785$). Встановлено, що серед моделей машинного навчання XGBoost показала найкращі результати для прогнозування зчеплення ($R^2 = 0.6764$, RMSE = 5.09 кПа) та UCS ($R^2 = 0.4427$, RMSE = 145.42 МПа), тоді як RF була оптимальною для прогнозування кута внутрішнього тертя ($R^2 = 0.4955$, RMSE = 2.35°) при цьому модель ANN стабільно поступалася деревоподібним моделям. Визначено, що лінійна регресія забезпечила конкурентоспроможні базові результати, що свідчить про переважно лінійний характер зв'язку між даними GR-каротажу та механічними властивостями при роздільній здатності каротажних вимірювань.

Наукова новизна. У роботі системно обґрунтовано використання даних гамма-каротажу як основного предиктора механічних властивостей гірських порід. Запропонований підхід рідко застосовувався в попередніх дослідженнях, де переважно використовувалися акустичні або густинні каротажні дані. Дослідження пропонує перше системне емпіричне порівняння лінійних і нелінійних моделей із використанням GR як основної вхідної ознаки, а також формує практичні регресійні рівняння, придатні для безпосереднього застосування в польових умовах.

Практична значимість. Отримані регресійні рівняння та перевірена схема моделювання дають змогу інженерам виконувати швидке геомеханічне оцінювання з використанням широко доступних даних гамма-каротажу, зменшуючи залежність від дорогих і тривалих лабораторних випробувань. Такий підхід є особливо цінним на ранніх стадіях геологорозвідувальних робіт, а також у гірничих умовах, де доступність кернового матеріалу є обмеженою.

Ключові слова: *гамма-каротаж; механічні параметри гірських порід; лінійна регресія; машинне навчання; геомеханіка*