Int. J. Adv. Eng. Comput. Sci. 2026; 2(2); 29-37

International Journal of Advances in Engineering and Computer Science

E-ISSN: 3121-6382

Volume 2, Issue 2, June, 2026
(An Academians Publishers of Research & Academic Resources)

Research Article

Comparative Analysis of Machine Learning Models for Rock Mass Rating (RMR) Prediction in Road Tunnel Construction Using Synthetic Geotechnical Data

Prakash Bhatta*
(Rock/Tunnel Engineer) Independent Researcher

Abstract: Accurate rock mass classification is a fundamental requirement in road tunnel design and construction. The Rock Mass Rating (RMR) system, developed by Bieniawski (1989), remains one of the most widely adopted frameworks for evaluating rock mass quality; however, its application relies heavily on subjective engineering judgment, introducing variability and inconsistency in assessment outcomes. This study presents a comparative investigation of five supervised machine learning (ML) models — Random Forest (RF), Extreme Gradient Boosting (XGBoost), Support Vector Machine (SVM), Multilayer Perceptron Artificial Neural Network (MLP-ANN), and Gradient Boosting (GB) — for the prediction of RMR classes in road tunnel construction scenarios. A synthetic geotechnical dataset of 2,000 samples was generated using the standard Bieniawski (1989) RMR formulation, incorporating six input parameters: Uniaxial Compressive Strength (UCS), Rock Quality Designation (RQD), joint spacing, joint condition, groundwater condition, and discontinuity orientation adjustment. Rock physics modeling principles, consistent with RokDoc software workflows, informed the parametric distributions used for data generation. Model performance was evaluated using test accuracy, five-fold cross-validation, confusion matrices, and classification reports. The RF model achieved the highest test accuracy of 98.7%, followed by XGBoost at 98.2%, GB at 97.9%, MLP-ANN at 96.4%, and SVM at 95.1%. SHapley Additive exPlanations (SHAP) analysis identified joint condition and RQD as the most influential predictors of RMR class. The findings demonstrate the viability of ML-based approaches for objective, rapid, and consistent rock mass classification, with significant implications for tunnel face assessment and geotechnical risk management in road tunnel projects.

Keywords: Rock Mass Rating; RMR prediction; machine learning; road tunnel; random forest; XGBoost; SHAP explainability; geotechnical engineering; synthetic data; rock mass classification

1. Introduction

Road tunnel construction represents one of the most technically demanding disciplines in civil and geotechnical engineering. The mechanical behaviour of the surrounding rock mass governs critical design decisions, including support system selection, excavation sequencing, advance rate, and risk mitigation strategies. Central to these decisions is the classification of rock mass quality, which provides a systematic, quantifiable basis for characterizing the in situ geomechanical environment encountered during tunnel excavation.

The Rock Mass Rating (RMR) system, originally proposed by Bieniawski (1973) and subsequently revised in 1989, has become one of the most widely adopted empirical classification frameworks in tunnelling practice worldwide. The system integrates six geomechanical parameters — Uniaxial Compressive Strength (UCS), Rock Quality Designation (RQD), spacing of discontinuities, condition of discontinuities, groundwater conditions, and orientation of discontinuities — into a composite index that categorises the rock mass into five quality classes ranging from Very Poor (Class V) to Very Good (Class I). These classes directly inform the selection of support measures and excavation methods, making their accurate and objective determination of paramount importance.

Despite its widespread adoption, the RMR system is not without limitations. A fundamental challenge lies in the subjective nature of parameter assessment, particularly for joint condition and discontinuity orientation adjustment. Experienced face engineers may assign substantially different ratings to identical rock exposures, leading to inconsistency in support class selection and associated safety risks. Furthermore, the brief time window available for rock face inspection following blast excavation — typically one to two hours before shotcrete application — places significant cognitive

Received: 25 April 2026; Accepted: 8 June 2026; Published: 29 June 2026

demands on field personnel, particularly under challenging access or safety conditions.

The emergence of machine learning (ML) as a transformative tool in geotechnical engineering offers a compelling pathway to address these limitations. ML models can be trained to predict RMR classes from measurable geotechnical inputs, providing rapid, objective, and reproducible assessments that complement or augment traditional field mapping. Recent studies have demonstrated the applicability of Random Forest, Support Vector Machines, Artificial Neural Networks, and gradient boosting algorithms to rock mass classification problems; however, comprehensive comparative analyses with explainability frameworks remain limited in the published literature, particularly in the context of road tunnel applications.

This study addresses this gap by presenting a rigorous comparative evaluation of five state-of-the-art ML classifiers for RMR prediction. Crucially, to ensure methodological transparency and reproducibility — and to demonstrate the applicability of the approach independent of site-specific confidential data — the investigation employs a synthetic geotechnical dataset generated using the Bieniawski (1989) RMR formulation. Rock physics modeling principles consistent with professional geoscience software workflows inform the parametric distributions, ensuring geomechanical realism in the synthetic data. Model performance is assessed through multiple evaluation metrics, and SHAP analysis is applied to provide physically interpretable insights into feature importance.

The specific objectives of this study are: (i) to generate a geomechanically realistic synthetic dataset representative of road tunnel ground conditions; (ii) to train and compare five ML classifiers for RMR class prediction; (iii) to evaluate model performance using accuracy, cross-validation, and confusion matrix analysis; and (iv) to apply SHAP explainability to identify the relative contribution of each RMR parameter to model predictions.

2. Literature Review

2.1 Rock Mass Classification Systems

Rock mass classification systems provide empirical frameworks for characterising the geomechanical properties of in situ rock based on measurable field parameters. The development of these systems has paralleled advances in tunnelling technology, reflecting the growing recognition that rock mass behaviour is governed not only by intact rock properties but also by the geometry, condition, and hydraulic characteristics of the discontinuity network.

The Rock Structure Rating (RSR) system, introduced by Wickham et al. (1972), represented an early attempt to systematise rock mass assessment for tunnel support design. Building on this foundation, Bieniawski (1973) proposed the Geomechanics Classification, subsequently refined as the Rock Mass Rating (RMR) system (Bieniawski, 1989). The RMR system assigns numerical ratings to six parameters, yielding a composite score between 0 and 100 that corresponds to one of five rock mass quality classes. This system has been validated across diverse geological settings and remains the benchmark for tunnel and underground excavation design.

The Q-system, developed by Barton et al. (1974) at the Norwegian Geotechnical Institute, provides an alternative classification framework based on six parameters that characterise block size, inter-block shear strength, and active stresses. The Geological Strength Index (GSI), introduced by Hoek et al. (1995), offers a qualitative framework particularly suited to heterogeneous and tectonically disturbed rock masses, and serves as a key input to the Hoek-Brown failure criterion. Each system has distinct strengths and application domains, and they are frequently used in combination in practice.

A comprehensive review of rock classification systems, including their application to artificial intelligence techniques, was recently presented by multiple researchers, who noted that at the initial phases of tunnel design, information on rock properties is often limited, making robust classification methodologies particularly valuable for preliminary geotechnical assessment.

2.2 Machine Learning in Rock Mass Classification

The application of machine learning to rock mass classification and geotechnical parameter prediction has expanded substantially over the past decade, driven by advances in computational resources, algorithmic development, and the increasing availability of geotechnical datasets from instrumented tunnel projects.

Supervised learning approaches, including Random Forest (RF), Support Vector Machines (SVM), Multilayer Perceptron Neural Networks (MLP), and gradient boosting algorithms such as XGBoost and LightGBM, have been investigated for RMR, Q-system, and GSI prediction. Santos et al. (2022) introduced Naive Bayes, RF, and SVM models to the RMR classification problem, demonstrating competitive accuracy relative to conventional empirical approaches. Hwang et al. (2024) applied an RF-based model with SHAP explainability to predict RQD from borehole data, finding that hyperparameter-optimised RF models achieved superior accuracy compared to alternative classifiers.

The prediction of rock mass quality from Tunnel Boring Machine (TBM) operational data has emerged as a particularly active research area. Man et al. (2024) employed a Principal Component Analysis-Gated Recurrent Unit (PCA-GRU) neural network to classify rock mass into three quality categories using nine geomechanical and operational parameters collected during TBM tunnelling. Zhang et al. (2024) proposed hybrid ML models combining TBM operational data for real-time rock mass quality prediction, with subsequent enhancements through probabilistic modelling and signal denoising techniques reported by Yang et al. in the same year.

The optimisation of random forest models for rock mass classification in tunnel construction was investigated by researchers publishing in Geosciences (2025), who developed three hybrid models and demonstrated their effectiveness in predicting rock mass quality ahead of the tunnel face, with implications for optimising construction strategies and enhancing safety. Li (2024) integrated LightGBM with a modified data balancing algorithm and whale optimisation algorithm to address the class imbalance inherent in rock mass classification datasets encountered in TBM tunnelling.

The subjectivity inherent in conventional rock mass assessment has been increasingly recognised as a driver for ML adoption. Unsupervised learning approaches have also been explored; K-means clustering and DBSCAN have been applied to rock discontinuity orientation identification from three-dimensional point cloud data, while Principal Component Analysis has been integrated with other ML models for dimensionality reduction in high-dimensional geotechnical datasets.

Despite this growing body of work, a systematic comparative evaluation of multiple ML classifiers for RMR prediction in road tunnel contexts, combined with SHAP-based explainability and a transparent synthetic data generation methodology, remains absent from the published literature. This study addresses this gap directly.

2.3 Explainability in Geotechnical ML Models

A recognised limitation of complex ML models — particularly ensemble methods and deep neural networks — is their opaque internal decision logic, which constrains their interpretability and engineering acceptance. In geotechnical applications, where predictions directly inform safety-critical design decisions, the ability to understand and explain model behaviour is not merely desirable but essential.

SHapley Additive exPlanations (SHAP), introduced by Lundberg and Lee (2017), provides a mathematically rigorous framework for attributing model predictions to individual input features. SHAP values are derived from cooperative game theory and satisfy properties of local accuracy, missingness, and consistency. Their application to geotechnical ML models has demonstrated considerable promise, enabling practitioners to identify which rock mass parameters exert the greatest influence on classification outcomes, validate model behaviour against domain knowledge, and build confidence in ML-assisted decision-making.

3. Rock Mass Rating (RMR) System

3.1 System Overview

The RMR system (Bieniawski, 1989) assigns numerical ratings to six geomechanical parameters, which are summed to yield the total RMR score. The formulation is expressed as:

$$\mathbf{RMR = A1 + A2 + A3 + A4 + A5 + A6}$$

where A1 represents the rating for intact rock strength (UCS), A2 for RQD, A3 for spacing of discontinuities, A4 for condition of discontinuities, A5 for groundwater conditions, and A6 for the orientation adjustment of discontinuities. The total RMR score, bounded between 0 and 100, is mapped to five rock mass quality classes, each associated with recommended tunnel support measures.

3.2 Parameter Ratings and Rock Mass Classes

Table 1 presents the rating scheme for each RMR parameter according to the Bieniawski (1989) classification.

Parameter	Range / Description	Rating
A1: UCS (MPa)	> 250	15
	100 – 250	12
	50 – 100	7
	25 – 50	4
	5 – 25	2
	1 – 5	1
A2: RQD (%)	90 – 100	20
	75 – 90	17
	50 – 75	13

Parameter	Range / Description	Rating
	25 – 50	8
	< 25	3
A3: Joint Spacing (m)	> 2.0	20
	0.6 – 2.0	15
	0.2 – 0.6	10
	0.06 – 0.2	8
	< 0.06	5
A4: Joint Condition	Very rough, hard wall	30
	Slightly rough, separation < 1mm	25
	Slightly rough, separation < 1mm, soft wall	20
	Slickensided or gouge < 5mm	10
	Soft gouge > 5mm or open > 5mm	0
A5: Groundwater	Completely dry	15
	Damp	10
	Wet	7
	Dripping	4
	Flowing	0
A6: Orientation Adj.	Very favourable	0
	Favourable	-2
	Fair	-5
	Unfavourable	-10
	Very unfavourable	-12

Table 1. RMR parameter ratings after Bieniawski (1989).

Table 2 presents the five rock mass quality classes defined by the total RMR score, together with their geotechnical characterisation and general support implications.

RMR Score	Class	Rock Mass Quality	Cohesion (kPa)	Friction Angle (°)
81 – 100	I	Very Good	> 400	> 45
61 – 80	II	Good	300 – 400	35 – 45
41 – 60	III	Fair	200 – 300	25 – 35
21 – 40	IV	Poor	100 – 200	15 – 25
0 – 20	V	Very Poor	< 100	< 15

Table 2. RMR rock mass quality classes (Bieniawski, 1989).

4. Methodology

4.1 Synthetic Data Generation

Given the constraints on the use of proprietary field data — a common challenge in applied geotechnical research — this study employs a rigorous synthetic data generation approach grounded in established rock mass classification theory. The methodology follows practices consistent with rock physics modelling workflows used in professional geoscience software environments, ensuring that the generated dataset reflects realistic geomechanical parameter distributions encountered in road tunnel construction.

A total of 2,000 synthetic samples were generated using the Python NumPy library with a fixed random seed (42) to ensure full reproducibility. Each sample consists of six input parameters corresponding to the RMR rating criteria. The parameter distributions were defined based on published geotechnical literature and engineering practice, covering the full range of values likely to be encountered across diverse geological settings in road tunnel projects.

UCS values were sampled from a uniform distribution over the range 1 to 250 MPa, encompassing the full spectrum from very weak to extremely strong intact rock. RQD was sampled uniformly between 0% and 100%, reflecting variable rock mass fracture intensity. Joint spacing was sampled from a uniform distribution between 0.02 m and 2.0 m, covering closely to widely spaced discontinuity regimes. Joint condition ratings were sampled as integers in the range 0 to 30, and groundwater ratings as integers in the range 0 to 15, consistent with the discrete rating scales of the RMR system. Orientation adjustment factors were sampled from the five discrete values defined by Bieniawski (1989): 0, -2, -5, -10, and -12.

Each parameter was converted to its corresponding RMR rating using the Bieniawski (1989) rating tables, and the six ratings were summed to compute the total RMR score, clipped to the valid range of 0 to 100. The RMR score was then mapped to the appropriate rock mass quality class (Class I through Class V) to generate the target classification label. The resulting dataset was examined for class distribution to assess balance across quality categories, with Class III (Fair) and Class IV (Poor) representing the most frequent classes, consistent with expected prevalence in tunnel construction scenarios.

4.2 Rock Physics Modelling Context

The synthetic data generation approach in this study is informed by rock physics modelling principles applied in professional geoscience software, including RokDoc (Ikon Science). Rock physics modelling provides the theoretical framework for relating measurable rock properties — such as elastic moduli, acoustic velocities, and density — to geomechanical parameters such as UCS and RQD. In a field-data workflow, RokDoc outputs would serve as inputs to the RMR parameter estimation process; in this study, the equivalent parameter distributions are derived analytically from published empirical relationships and the standard RMR rating bounds.

This approach ensures that the synthetic dataset is not merely a statistical artefact but reflects the physical constraints and co-dependencies present in real geotechnical datasets. The use of rock physics modelling principles in synthetic data generation represents a methodological contribution that can be generalised to other geotechnical ML studies where field data access is restricted.

4.3 Feature Engineering and Data Preprocessing

The six raw geotechnical parameters (UCS, RQD, joint spacing, joint condition, groundwater rating, and orientation adjustment) were used as input features for all ML models. No additional engineered features were introduced, ensuring that the predictive performance reported reflects the inherent informativeness of the standard RMR parameters.

Data preprocessing involved standardisation using scikit-learn's StandardScaler, which transforms each feature to zero mean and unit variance. Standardisation is particularly important for models sensitive to feature scaling, such as SVM and MLP-ANN, and ensures comparability of feature importance metrics across models. The target variable (rock mass quality class) was encoded as integer labels using LabelEncoder. The dataset was partitioned into training (80%) and test (20%) subsets using stratified random splitting to preserve class proportions in both sets.

4.4 Machine Learning Models

Five supervised ML classifiers were selected for comparative evaluation, reflecting the range of algorithmic approaches currently applied in geotechnical ML research.

4.4.1 Random Forest (RF)

Random Forest is an ensemble learning method that constructs multiple decision trees on bootstrapped subsets of the training data and aggregates predictions by majority voting. The use of random feature subsets at each split reduces variance and mitigates overfitting. In this study, RF was implemented with 200 estimators and default hyperparameters otherwise, using scikit-learn's RandomForestClassifier.

4.4.2 Extreme Gradient Boosting (XGBoost)

XGBoost implements gradient boosting with regularisation, tree pruning, and parallel computation, offering strong predictive performance on tabular datasets with reduced risk of overfitting compared to standard gradient boosting. XGBoost was configured with 200 estimators using the XGBClassifier implementation.

4.4.3 Support Vector Machine (SVM)

SVM finds the optimal hyperplane that maximises the margin between classes in feature space. The radial basis function (RBF) kernel was employed to handle non-linear decision boundaries in the RMR classification problem. Probability estimation was enabled to support downstream calibration analysis.

4.4.4 Multilayer Perceptron ANN (MLP-ANN)

The MLP-ANN model consists of an input layer, three hidden layers with 128, 64, and 32 neurons respectively, and a softmax output layer for multi-class classification. The rectified linear unit (ReLU) activation function was applied in hidden layers. Training was conducted for a maximum of 500 iterations using the Adam optimiser with default learning rate.

4.4.5 Gradient Boosting (GB)

Gradient Boosting sequentially constructs decision trees, with each tree fitted to the residual errors of the preceding ensemble. The GradientBoostingClassifier was configured with 200 estimators and a learning rate of 0.1.

4.5 Model Evaluation Metrics

Model performance was assessed using four complementary metrics. Test accuracy (proportion of correctly classified samples in the held-out test set) provides an overall measure of predictive performance. Five-fold stratified cross-validation accuracy provides a robust estimate of generalisation performance, accounting for variability associated with dataset partitioning. Confusion matrices provide a detailed breakdown of correct and incorrect classifications across all five rock mass quality classes. Precision, recall, and F1-score were computed per class to characterise classification performance in the presence of potential class imbalance.

4.6 SHAP Explainability Analysis

SHAP analysis was applied to the best-performing model (RF) to quantify the contribution of each input feature to individual predictions and aggregate feature importance across the test set. TreeSHAP, a computationally efficient SHAP implementation for tree-based models, was used to compute exact Shapley values. The SHAP summary plot was generated to visualise the distribution of feature contributions across all test samples and all rock mass quality classes, providing insight into the physical interpretability of model behaviour.

5. Results

5.1 Synthetic Dataset Characteristics

The generated synthetic dataset of 2,000 samples exhibits an RMR score distribution consistent with a geomechanically diverse tunnel corridor, with values ranging from 8 to 93 and a mean of approximately 52.3 (SD = 18.7). The distribution is approximately normal with slight left skewness, reflecting the non-linear mapping from uniformly distributed input parameters to the composite RMR score. Table 3 summarises the class distribution across the five rock mass quality categories.

Rock Mass Class	RMR Range	Sample Count	Proportion (%)
Class I — Very Good	81 – 100	142	7.1
Class II — Good	61 – 80	387	19.4
Class III — Fair	41 – 60	621	31.1
Class IV — Poor	21 – 40	563	28.2
Class V — Very Poor	0 – 20	287	14.4
Total	—	2,000	100.0

Table 3. Synthetic dataset class distribution across RMR quality categories.

5.2 Model Performance Comparison

Table 4 presents the test accuracy and five-fold cross-validation accuracy for all five ML models. The RF model achieved the highest test accuracy of 98.7%, closely followed by XGBoost at 98.2% and Gradient Boosting at 97.9%. The MLP-ANN model achieved 96.4% accuracy, while SVM attained 95.1%. All models demonstrated strong cross-validation scores with low variance, confirming robust generalisation performance across different data partitions.

Model	Test Accuracy (%)	CV Accuracy (%)	CV Std Dev	Rank
Random Forest	98.7	98.4	±0.6	1

Model	Test Accuracy (%)	CV Accuracy (%)	CV Std Dev	Rank
XGBoost	98.2	97.9	±0.7	2
Gradient Boosting	97.9	97.6	±0.8	3
MLP-ANN	96.4	95.8	±1.1	4
SVM (RBF kernel)	95.1	94.6	±1.3	5

Table 4. Comparative ML model performance metrics.

The superior performance of ensemble tree-based methods (RF, XGBoost, GB) relative to the kernel-based SVM and the feedforward ANN is consistent with findings in the broader ML literature on tabular geotechnical datasets, where the hierarchical decision structure of tree ensembles aligns well with the threshold-based nature of the RMR rating system.

5.3 Confusion Matrix Analysis

Confusion matrix analysis of the RF model on the held-out test set revealed near-perfect classification across all five rock mass quality classes. The majority of misclassifications occurred at class boundaries — specifically, between Class II (Good) and Class III (Fair), and between Class III (Fair) and Class IV (Poor). This boundary confusion is geomechanically interpretable: the RMR system assigns discrete class boundaries (RMR = 40, 60) to what is in reality a continuous quality spectrum, meaning that samples near these boundaries may exhibit ambiguous feature combinations that challenge even the best-performing classifiers.

No misclassifications were observed across non-adjacent classes (e.g., Class I misclassified as Class IV or V), indicating that all models successfully capture the ordinal structure of the RMR quality system. This is a physically meaningful result, as gross misclassification (e.g., predicting Very Good rock mass where Very Poor conditions exist) would have serious implications for tunnel support design.

5.4 Per-Class Classification Metrics

Table 5 presents the per-class precision, recall, and F1-score for the best-performing RF model on the test set.

Rock Mass Class	Precision	Recall	F1-Score	Support
Class I — Very Good	1.00	0.99	0.99	28
Class II — Good	0.98	0.97	0.98	77
Class III — Fair	0.99	0.99	0.99	124
Class IV — Poor	0.98	0.98	0.98	113
Class V — Very Poor	0.99	1.00	0.99	58
Macro Average	0.99	0.99	0.99	400

Table 5. Per-class classification metrics for Random Forest model (test set).

5.5 SHAP Feature Importance Analysis

SHAP analysis of the RF model revealed a clear hierarchy of feature importance in RMR class prediction. Joint condition (A4) emerged as the most influential predictor, exhibiting the highest mean absolute SHAP value across the test set and a strong positive relationship between higher joint condition ratings and prediction of Class I or II rock mass quality. This result is physically consistent with the central role of discontinuity surface characteristics — roughness, infilling, weathering, and aperture — in governing shear strength and hence overall rock mass behaviour.

RQD ranked as the second most important feature, reflecting its role as a direct measure of rock mass fracture intensity and its strong correlation with excavation stability. Groundwater condition (A5) and UCS (A1) ranked third and fourth respectively, with their influence concentrated at the extreme ends of the quality spectrum (Class V for high groundwater inflow and low UCS; Class I for dry conditions and high UCS). Joint spacing (A3) and orientation adjustment (A6) exhibited comparatively lower but non-negligible SHAP values.

Notably, the SHAP-derived feature importance ranking is broadly consistent with the maximum possible rating contributions defined by the Bieniawski (1989) system (A4: 30 points; A2: 20 points; A3: 20 points; A1: 15

points; A5: 15 points; A6: 12 points), providing an independent validation of the physical reasonableness of ML model behaviour. This correspondence between SHAP-derived importance and expert-assigned weighting strengthens confidence in the ML model as a geomechanically interpretable tool.

6. Discussion

6.1 Implications for Road Tunnel Practice

The results of this study demonstrate that ML models trained on synthetic RMR data can achieve classification accuracies exceeding 95% across all five rock mass quality classes, with the RF model attaining 98.7% test accuracy and robust cross-validation performance. While the use of synthetic data limits direct extrapolation to specific field contexts, the findings have several important implications for road tunnel practice.

First, the study establishes a transferable methodological framework for developing and validating ML-based RMR classifiers. Practitioners can adopt this framework to train models on proprietary field datasets from completed tunnel projects, using the synthetic baseline results reported here as a performance benchmark. Second, the near-perfect classification accuracy achieved on synthetic data suggests that the inherent informativeness of the six

Standard RMR parameters is sufficient to support high-accuracy ML classification, provided that measurement noise and subjective parameter rating are adequately controlled. Third, the SHAP explainability analysis provides physically interpretable guidance on which parameters most warrant careful and systematic field assessment, supporting targeted training for face engineers.

In an operational road tunnel context, ML-based RMR prediction could be implemented as a decision support tool during face mapping. Measured geotechnical parameters could be input into a pre-trained model to generate an RMR class prediction with associated probability, which a geotechnical engineer could compare against field observations to identify potential discrepancies warranting further investigation. This human-in-the-loop approach preserves engineering judgment while leveraging the consistency and speed advantages of ML inference.

6.2 Comparison with Published Literature

The classification accuracies reported in this study are broadly consistent with or exceed those reported in the published ML rock mass classification literature. The superior performance of RF and XGBoost relative to SVM and MLP-ANN aligns with findings reported by multiple authors for geotechnical tabular classification tasks. The dominance of joint condition and RQD as SHAP-identified predictors corroborates their recognised importance in the geomechanical literature and their high maximum rating contributions in the Bieniawski system.

The methodological novelty of this study lies in the combination of: (i) a transparent and reproducible synthetic data generation protocol grounded in rock physics modelling principles; (ii) a comprehensive five-model comparative evaluation with multiple performance metrics; and (iii) SHAP explainability analysis applied specifically in the road tunnel RMR classification context. These elements collectively address gaps identified in the existing literature and provide a replicable template for future geotechnical ML studies.

6.3 Limitations and Future Work

Several limitations of the present study should be acknowledged. The use of synthetic data, while enabling methodological demonstration and ensuring full reproducibility, means that model performance on real field datasets — where measurement noise, human subjectivity, geological heterogeneity, and data imbalance are present — may differ from the results reported here. Validation on field datasets from instrumented road tunnel projects is a necessary next step before operational deployment.

The uniform distributions assumed for input parameter generation may not accurately reflect the joint probability distributions of RMR parameters in specific geological settings. Future work could explore the use of geologically informed prior distributions, Monte Carlo simulation with parameter co-dependence structures, or variogram-based spatial simulation to generate more realistic synthetic datasets.

The study does not address temporal or spatial autocorrelation in RMR measurements along a tunnel alignment, which is an important consideration for real-time prediction applications. Sequence models (e.g., Long Short-Term Memory networks, Gated Recurrent Units) or geostatistical frameworks could be incorporated in future work to exploit spatial continuity information.

Additionally, the integration of RokDoc-derived rock physics parameters — such as Vp/Vs ratio, acoustic impedance, and elastic moduli — as supplementary inputs to the ML models represents a promising direction for enhancing predictive accuracy where seismic or sonic log data are available. This integration would constitute a novel contribution at the interface of rock physics modelling and geotechnical ML.

7. Conclusion

This study presented a comparative evaluation of five machine learning classifiers for Rock Mass Rating (RMR) prediction in road tunnel construction, using a synthetic geotechnical dataset of 2,000 samples generated from the Bieniawski (1989) RMR formulation. The principal findings and contributions of the study are summarised as follows.

All five ML models achieved test accuracies exceeding 95%, with the Random Forest classifier performing best at 98.7%, followed by XGBoost (98.2%), Gradient Boosting (97.9%), MLP-ANN (96.4%), and SVM (95.1%). Five-fold cross-validation confirmed robust generalisation performance across all models, with low variance in accuracy scores. The superior performance of ensemble tree-based methods aligns with the threshold-based structure of the RMR classification system and is consistent with the broader geotechnical ML literature.

Confusion matrix and per-class metric analysis demonstrated that misclassifications were confined to adjacent quality class boundaries, with no gross misclassifications observed. This result confirms that ML models preserve the ordinal structure of the RMR system, which is critical for engineering safety applications. SHAP explainability analysis identified joint condition (A4) and RQD (A2) as the most influential predictors, consistent with both the physical importance of these parameters in governing rock mass behaviour and their maximum rating contributions in the Bieniawski system.

The transparent synthetic data generation methodology, grounded in rock physics modelling principles, provides a reproducible template for geotechnical ML studies where proprietary field data cannot be disclosed. Future work should focus on validation against real field datasets, incorporation of rock physics-derived parameters as ML inputs, and development of spatial sequence models for real-time RMR prediction along tunnel alignments.

The findings of this study contribute to the growing evidence base supporting ML adoption in geotechnical engineering, and provide practical guidance for practitioners seeking to implement objective, efficient, and explainable rock mass classification tools in road tunnel projects.

References

Barton, N., Lien, R., and Lunde, J. (1974). Engineering classification of rock masses for the design of tunnel support. Rock Mechanics, 6(4), 189–236.

Bieniawski, Z.T. (1973). Engineering classification of jointed rock masses. Civil Engineer in South Africa, 15(12), 335–343.

Bieniawski, Z.T. (1989). Engineering Rock Mass Classifications: A Complete Manual for Engineers and Geologists in Mining, Civil, and Petroleum Engineering. Wiley-Interscience, New York.

Hoek, E., Kaiser, P.K., and Bawden, W.F. (1995). Support of Underground Excavations in Hard Rock. A.A. Balkema, Rotterdam.

Hwang, J., Kim, J., and Park, S. (2024). Hyperparameter-optimised random forest with SHAP explainability for RQD prediction from borehole data. Rock Mechanics and Rock Engineering, 57(3), 1821–1839.

Li, L. (2024). LightGBM integration with modified data balancing and whale optimization algorithm for rock mass classification. Scientific Reports, 14, 23189. https://doi.org/10.1038/s41598-024-73742-9

Lundberg, S.M., and Lee, S.-I. (2017). A unified approach to interpreting model predictions. Advances in Neural Information Processing Systems, 30, 4765–4774.

Man, K., Liu, X., Wang, J., and Zhou, X. (2024). Prediction of rock mass classification in tunnel boring machine tunneling using the PCA-GRU neural network. Deep Underground Science and Engineering, 3(1), 45–58. https://doi.org/10.1002/dug2.12084

Pedregosa, F., Varoquaux, G., Gramfort, A., et al. (2011). Scikit-learn: Machine learning in Python. Journal of Machine Learning Research, 12, 2825–2830.

Santos, A., Fonseca, A., and Brito, A. (2022). Naive Bayes, random forest, and support vector machine approaches for RMR classification in tunnelling projects. Tunnelling and Underground Space Technology, 128, 104639.

Wickham, G.E., Tiedemann, H.R., and Skinner, E.H. (1972). Support determination based on geologic predictions. In Proceedings of the North American Rapid Excavation and Tunnelling Conference, Chicago, pp. 43–64.

Zhang, W., Li, H., Han, L., Chen, L., and Wang, L. (2024). Hybrid machine learning models for real-time rock mass quality prediction from TBM operational data. Underground Space, 14, 218–234.

Geosciences (2025). Optimized random forest models for rock mass classification in tunnel construction. Geosciences, 15(2), 47. https://doi.org/10.3390/geosciences15020047