



OPEN

# Prediction of jumbo drill penetration rate in underground mines using various machine learning approaches and traditional models

Sasan Heydari<sup>1</sup>, Seyed Hadi Hoseinie<sup>1</sup>✉ & Raheb Bagherpour<sup>1</sup>

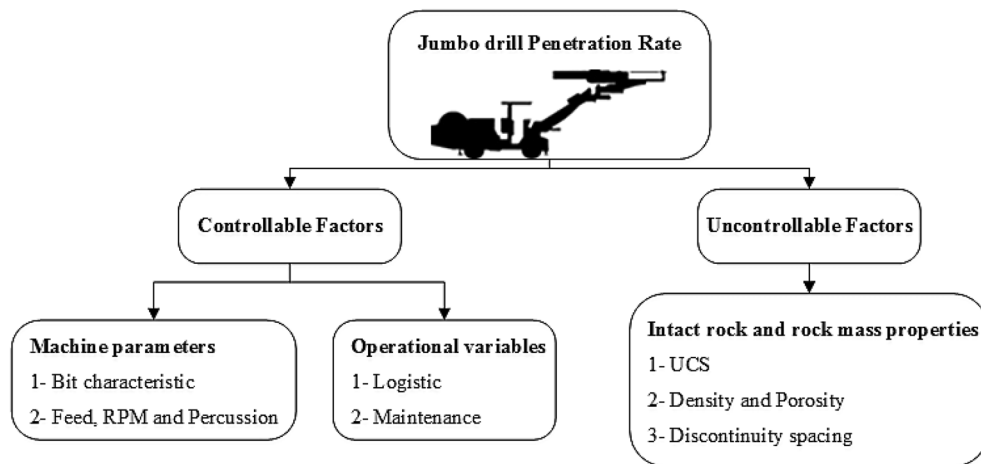
Estimating penetration rates of Jumbo drills is crucial for optimizing underground mining drilling processes, aiming to reduce costs and time. This study investigates various regression and machine learning methods, including Multilayer Perceptron (MLP), Support Vector Regression (SVR), and Random Forests (RF), to predict the penetration rates (ROP) using multivariate inputs such as operation parameters and rock mass characteristics. The Rock Mass Drillability Index (RDi), incorporating both intact rock properties and structural parameters, was utilized to characterize the rock mass. The dataset was split into 80% for training and 20% for testing. Performance metrics including correlation coefficient ( $R^2$ ), variance accounted for (VAF), mean absolute error (MAE), mean absolute percentage error (MAPE), and root mean square error (RMSE) were calculated for each method to evaluate the accuracy of the predictions. SVR exhibited the best prediction performance for ROP, achieving the highest  $R^2$ , lowest RMSE, MAE, and MAPE, as well as the largest VAF values of 0.94, 0.15, 0.11, 4.84, and 94.13 during training, and 0.91, 0.19, 0.13, 6.02, and 91.11 during testing, respectively. With this high accuracy, we conclude that the proposed machine learning algorithms are valuable and efficient predictors for estimating jumbo drill penetration rates in underground mining operations.

**Keywords** Penetration rate prediction, The Rock Mass Drillability Index (RDi), Traditional models, Multilayer perceptron neural networks (MLP), Support Vector Regression (SVR), Random Forests (RF)

The drill and blast method is the most important method used in underground excavation. In the drill and blast method in tunneling, drilling constitutes the largest cost and time<sup>1</sup>. In drilling operations, many factors such as geological, geotechnical, operational, and machine characteristics, affect drilling performance. These parameters can be generally classified into two major groups controllable and uncontrollable parameters<sup>2,3</sup>. Operational factors and machine characteristics are variables that can be controlled but geological and geotechnical conditions are unique to each site and cannot be easily altered<sup>4</sup>. The key factors that impact the Rate of Penetration (ROP) are shown in Fig. 1.

Predicting the ROP based on drilling variables is essential aiming to maximize ROP or minimize total time or cost. For that, the accuracy of the ROP model is crucial<sup>4</sup>. Despite the importance of predicting ROP for better drilling efficiency, accurately establishing a prediction model is challenging<sup>5</sup>. The empirical approach, commonly used to study penetration rates, utilizes field data and is developed for varying ground conditions. Various studies conducted by researchers on percussive and rotary drilling have revealed that the rate of penetration is contingent upon the properties of the rock<sup>6</sup>. The relationship between rock properties and ROP is complex and nonlinear. Many researchers developed statistical models to predict ROP based on the experimental data<sup>7–17</sup>. Various statistical models and classification systems have been developed to predict penetration rate, but they often lack generalizability across different drilling conditions<sup>18</sup>. Existing methods may not accurately predict ROP in different geological settings or with different drilling rigs<sup>19–24</sup>. Many of these statistical relations only show the effect of different rock parameters on ROP separately. Few of them can predict ROP based on few

Department of Mining Engineering, Isfahan University of Technology, Isfahan 84156-83111, Iran. ✉email: hadi.hoseinie@iut.ac.ir



**Figure 1.** The most important factors influencing drilling performance.

rock properties. The Norwegian Institute of Technology introduced the Drilling Rate Index (DRI) for assessing drillability in percussive drilling<sup>25</sup>. Linear predictive equations were developed correlating rate of penetration with various rock properties such as compressive strength, tensile strength, and Young's modulus. Despite investigating numerous rock mechanical properties, researchers found that only compressive strength, tensile strength, and Young's modulus were indicative parameters for predicting penetration rate<sup>26</sup>. Furthermore, researchers noted that penetration rate closely relates to bulk density, compressive strength, apparent porosity, P-wave velocity, and Schmidt hammer value<sup>27</sup>. Correlations between ROP and different rock hardness test results were demonstrated, with compressive strength and tensile strength showing a strong correlation with ROP<sup>9,16</sup>. Multiple regression analysis was employed to construct PR models for various drill types, including rotary, Down-the-Hole (DTH), and hydraulic top hammer drills<sup>17</sup>. Moreover, a new index named the Rock Penetrability Index (RPi) was developed, considering factors like UCS, Schimazek's F-abrasivity, Mohs hardness, rock texture, grain size, and Young's modulus<sup>28,29</sup>. Density was also identified as an important parameter for predicting ROP<sup>12</sup>. Apart from intrinsic rock properties, discontinuities, such as the Rock Quality Designation (RQD), significantly affect the cutting performance<sup>30</sup>. Additionally, researchers have presented a drillability prediction model based on parameters such as UCS, Schmidt hammer hardness value, quartz content, fragment size, alternation, joint dip, bit rotational speed, and thrust<sup>31</sup>. Furthermore, studies have explored the relationship between ROP and drill rig operational parameters, with theoretical and experimental investigations suggesting that ROP can be predicted by specific energy and certain design and operational parameters of the drilling machine<sup>32</sup>. Relationships between ROP and drilling thrust, as well as the influence of rock strength, drill rig power, drill bit shape, and geological discontinuities on ROP, have also been highlighted<sup>33</sup>. Some of the research findings include empirical models based on the uniaxial compressive strength of rocks, the impact of geological features on penetration rate, and the use of Artificial intelligence techniques to predict rock classification around tunnels. The studies also cover topics related to percussive and rotary drilling, the influence of different parameters on penetration rate, and the development of new models for predicting drilling performance<sup>34–51</sup>.

Statistical approaches in ROP prediction suggest relationships between input and output parameters, but these methods may not always effectively handle non-linear and complex problems<sup>52,53</sup>. The traditional multiple regression model lacks adequate prediction stability and struggles to solve nonlinear problems due to the impact of multicollinearity among independent variables<sup>54</sup>. Artificial Intelligence Based models have been introduced to address these limitations and, over time, have demonstrated improvements in prediction accuracy<sup>55</sup>.

The advancement of information technology has led to the emergence of intelligent drilling and completion technology in the field. Currently, there is a growing trend in the application of various machine learning techniques for estimating ROP as they continue to evolve and advance rapidly<sup>56</sup>. Due to the need for frequent recalibration with traditional physics-based models based on auxiliary data, machine learning models strive to overcome these challenges by utilizing data to identify correlations among various drilling variables<sup>57</sup>. Machine learning algorithms can take any number of measured variables as inputs, making them a powerful tool for ROP modeling<sup>4</sup>. ML algorithms have an advantage over analytical ROP models as they offer flexibility in model form, which allows them to effectively segment the drilling operational parameter space. Unlike analytical models, ML algorithms do not require predefined equations, as hyperparameters specific to each algorithm can control the model architecture. However, increased model complexity can lead to reduced interpretability and an increased risk of overfitting. Also, in the presence of multiple collinearity between input variables, Methods like neural networks often encounter challenges such as an unstable learning process and slow convergence speed in calculations<sup>58</sup>.

The application of machine learning techniques in underground mines extends beyond the study of ROP (Rate of Penetration), and nowadays, it is also utilized in other fields of engineering. With the widespread use of jumbodrills equipped with Measurement While Drilling (MWD) systems, it has become customary to employ machine learning and deep learning techniques for classifying the rock mass in tunnels, Predictive modeling of

drilling rate, as well as detecting deviations in bore holes within the face<sup>51</sup>. In addition to the aforementioned applications, given the paramount importance of safety in underground mines, the utilization of these techniques in various areas such as fire prediction, column stability, and rockburst in underground engineering structures is rapidly expanding<sup>59–61</sup>.

One of the early applications of Machine Learning for predicting drilling parameters, was presented by Arehart 1990<sup>62</sup>. Arehart employed Artificial neural networks (ANN) to predict a crucial drill bit parameter, specifically bit wear. Subsequently, Bilgesu et al. published the first study applying ANN for predicting the ROP<sup>63</sup>. In drilling, machine learning techniques such as ANN, random forests, and other regression and classification methods, along with deep learning methods, have demonstrated numerous applications. They have proven to be advantageous<sup>64</sup>. The findings suggest that artificial neural networks are the most commonly used machine learning technique for managing the rate of penetration. Among ANNs, basic models outperform modified versions in this context. However, while modified ANNs demonstrate greater accuracy in predicting ROP, they are not superior to other machine learning methods such as linear regression (LR) and random forest (RF) in making highly accurate predictions. These alternative approaches have been shown to be effective and practical in compensating for the limitations of ANNs in ROP management<sup>3,65</sup>. Moreover, similar studies have been conducted to predict drilling rate using machine learning techniques such as random forest. However, these models encounter constraints due to the utilization of laboratory samples. It is imperative that these models undergo testing in the field with real-time drilling data to gain further insight into their performance in practical scenarios<sup>66,67</sup>.

After reviewing the literature on the subject, it was discovered that only a few researchers have developed relationships to predict the penetration rate of jumbodrills. Kahraman calculated the performance prediction of a jumbo drill based on intact rock and rock mass properties<sup>45</sup>. The relationship he presented focuses on the discontinuities within the rock mass. Shen found relationships between the drilling rate and drilling parameters based on drilling data. To explore the variations in drilling parameters, drilling tests were conducted on a rock block. His relations do not take into account the characteristics of the rock mass<sup>68</sup>. Su developed multiple regression models using UCS, DRI, impact energy, blows per minute of the piston, hole area, and some rock properties for predicting the penetration rate and specific energy of drilling<sup>69</sup>.

Prasad developed an experimental methodology to measure the expected bit life along with ROP using a single test method and Compendious Index for Drillability<sup>70</sup>.

As far as the available literature suggests, there have been no attempts made to conduct a comprehensive review solely focusing on the recent advancements of machine learning techniques for predicting the jumbo drill penetration rate. Thus, the main aim of this paper is to fill this research gap and address crucial queries related to intelligent prediction techniques for jumbo drill's ROP. In this paper, in addition to regression methods, three machine learning methods, Multilayer perceptron (MLP), Support Vector Regression (SVR), and Random Forests (RF), are used to estimate ROP. All calculations in this paper were performed using a custom program code written in the R Statistical Programming Language.

## Case study and data collection

The objective of this paper is to develop an empirical model with machine learning algorithms that can predict the penetration rate of jumbo drills in underground mines. The collection of real field data constitutes a significant and essential part of this study. The models were developed using real data collected from seven underground lead and zinc mines that were extracted by sublevel stoping mining method located in the Irankouh and Sormeh mining districts of Iran. To create the models, the study considered six parameters, which relate to drilling and rock properties. In these mines, the tunnel face area is about 16–24 square meters and contains approximately 30 ~ 35 blasting holes per face. The jumbo drills used for the study included four different brands, with drill bit diameters of 51 and 64 mm. The collected parameters are described in Table 1.

## Relationship between rock mass parameters and ROP

Having reviewed the studies conducted thus far, it becomes apparent that several rock mass parameters significantly influence drilling operations. These include the origin of the rock formation, Mohs hardness, texture (grain shape and size), porosity, density, abrasiveness, rigidity, P-wave velocity, elasticity, plasticity, UCS (point load index and Schmidt hammer), tensile strength, and structural parameters (such as joints, cracks, and bedding) alongside RQD.

Parameter	Unit	Symbol	Description
Feed pressure	bar	$P_F$	The hydraulic pressure inside the cylinders required to keep the bit in contact with the bottom of the hole
Rotation pressure	bar	$P_R$	The pressure of the bit against the rock to maintain the required rotation
Hammer pressure	bar	$P_P$	The measurement of the impact pressure of the bit against the rock mass
Borehole Diameter	mm	$D_h$	Borehole Diameter, The diameter of the bits used to drill the boreholes is 51 and 64 mm
Rock mass drillability index	–	RDi	Rock mass drillability index, include physical parameters of the rock materia, strength parameters of the rock material and structural parameters of the rock mass
The rate of penetration	m/min	ROP	The result of dividing the drilling time by the length of the borehole

**Table 1.** Description of parameter symbols and values used in the model.

Rock classification poses a challenge due to the selection of parameters deemed most significant. No single parameter or index can comprehensively describe a jointed rock mass for engineering purposes. Consequently, it is impractical to incorporate all parameters into a classification system. In the rock mass drillability index classification scheme (RDi), four guiding principles have been adopted: a limited number of parameters, avoidance of equivalent parameters, categorization of parameters into groups, and field applicability of the classification. Adhering to these principles, the RDi assessment focuses on three categories of parameters: physical parameters of the rock material: texture and grain size, (b) strength parameters of the rock material: UCS and Mohs hardness, and (c) structural parameters of the rock mass: joint spacing, joint filling and aperture, and joint dipping. Except for UCS, all parameters are readily measurable in the field. Since it is not possible to measure the UCS in the field, one can use point load index and Schmidt hammer results as equivalent values for the UCS<sup>27</sup>.

Based on the points mentioned, there are many uncontrollable parameters that affect the penetration rate. However, incorporating all these variables is complex, time-consuming, and costly. Therefore, this paper employs the Rock Mass Drilling Index (RDi) as a representative of rock mass properties. RDi encompasses both intact rock properties and structural parameters, serving as a comprehensive indicator of rock mass characteristics.

To date, there has been no index used for underground mining that comprehensively encompasses all rock mass characteristics. The RDi, developed by Hosseini, provides a qualitative representation of the drilling rate by considering both intact rock properties and structural parameters of the rock mass<sup>27</sup>. This classification system is presented in Table 2, and the qualitative predictions are shown in Table 3. By summing the ratings in Table 2, a qualitative prediction of the rock mass drilling rate can be calculated.

Before commencing tunnel drilling, relevant information such as depth, cross-sectional area, water presence, and other parameters was gathered. Prior to jumbo drill operations in each tunnel face, data on discontinuity properties were collected, comprising joint spacing, joint aperture and filling, and the angle between the joint and borehole axis. A select number of suitable samples from each face underwent laboratory studies to determine the physical and mechanical properties of the rock. A total of 737 boreholes in 26 faces, equivalent to more than 2400 m, were considered for this study. Uniaxial compressive strength was calculated following International Society for Rock Mechanics (ISRM) guidelines (RTH 325–89), and density and porosity tests were conducted in accordance with ISRM Suggested Methods. Additional assessments included grain size, texture, and Mohs hardness determination. Each face was divided into distinct zones based on the rock mass drillability index (Refer to Fig. 2), and these zones were designated for borehole drilling. Data related to drilling in each specific area were recorded.

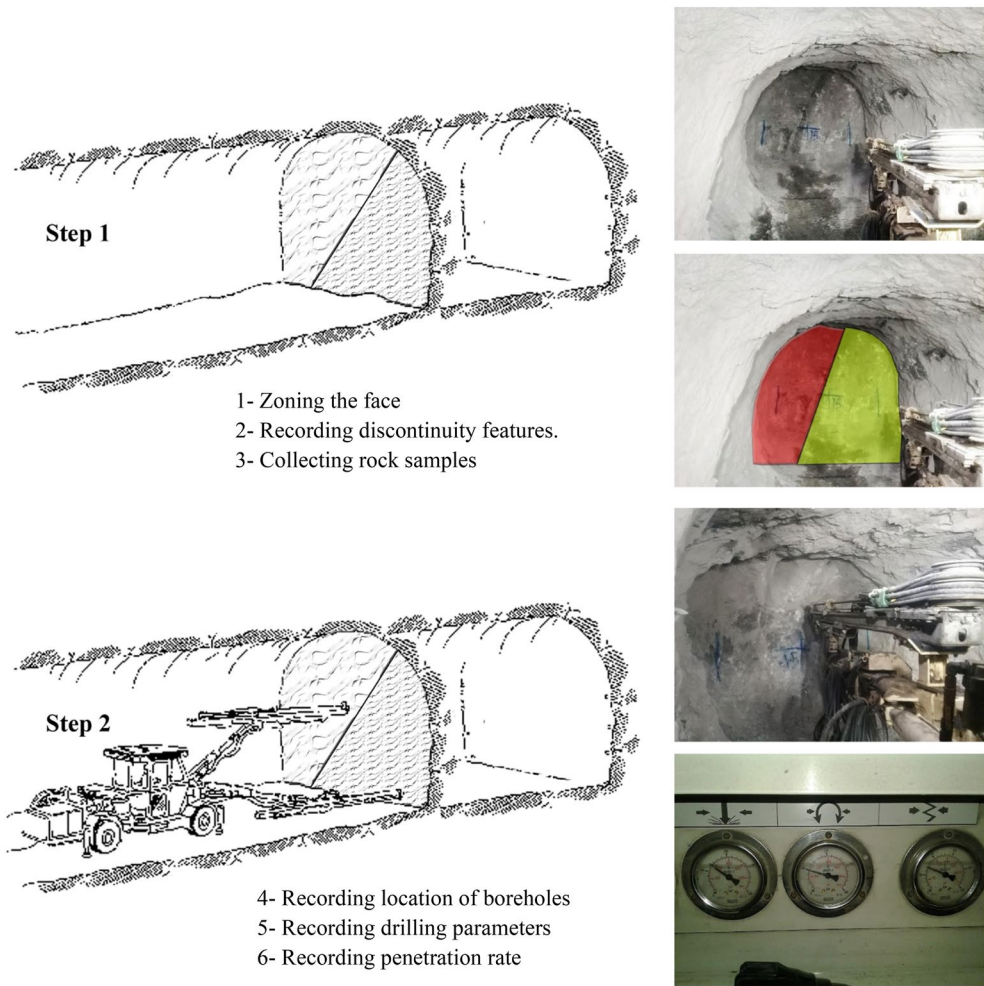
The Rock Mass Drillability Index (RDi) parameters collected from tunnels were rated as per Table 2 to classify rock mass drillability (see Table 3). To evaluate the correlation between RDi and the Rate of Penetration (ROP), correlation diagrams were created for two borehole diameters (51 and 64 mm) with respect to penetration rate (Fig. 3). The analysis revealed a linear increase in penetration rate with higher RDi values in the rock mass. This linear relationship was chosen based on the highest correlation coefficients among linear, logarithmic, power, and exponential functions. The variability in penetration rate for each RDi value is attributed to differences in machine operating parameters. To clarify the correlation between these two factors, the data related to each RDi

Texture	Porous	Fragmental	Granitoid	Porphyritic	Dense
Grain size	–	> 5 mm	2–5 mm	0.05–1 & 2–5 mm	0.05–1 mm
Rating	15	10	7	4	1
Mohs Hardness	1–3	3–4.5	4.5–6	6–7	> 7
Description	Very soft–soft	Comparatively soft	Comparatively hard	Hard	Very hard
Rating	18	13	9	4	1
UCS (MPa)	1–25	25–50	50–100	100–200	> 200
Description	Very low strength	Low strength	Average strength	High strength	Very high strength
Rating	22	16	11	6	2
Joints Spacing	> 2 m	1–2 m	0.5–1 m	0.15–0.5 m	0–0.15 m
Rating	18	13	9	5	1
Joint aperture & filling	Closed joint 0–2 mm	> 20 mm	12–20 mm	9–12 mm	2–9 mm
Rating	15	10	7	4	1
The angle between the joint & borehole axis	70°–90°	55°–70°	35°–55°	20°–35°	0°–20°
Rating	12	8	6	3	1

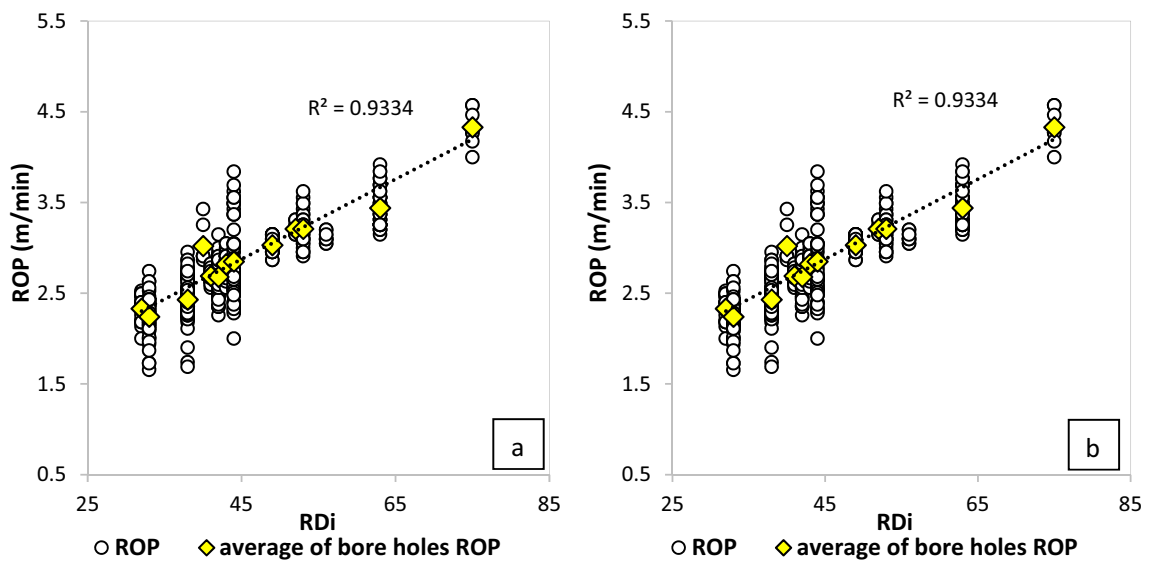
**Table 2.** Rock mass drillability index (RDi) classification<sup>27</sup>.

RDi	7–20	20–40	40–60	60–80	80–100
Prediction of drilling rate	Slow	Slow–medium	Medium	Medium–fast	Fast

**Table 3.** Qualitative prediction of the penetration rate of drilling in the rock mass using RDi<sup>27</sup>.



**Figure 2.** Different steps of data collection in this paper.



**Figure 3.** Relationship between RDi and all available datasets penetration rate and the average of many numbers of blasting holes (a) 51 mm (b) 64 mm.



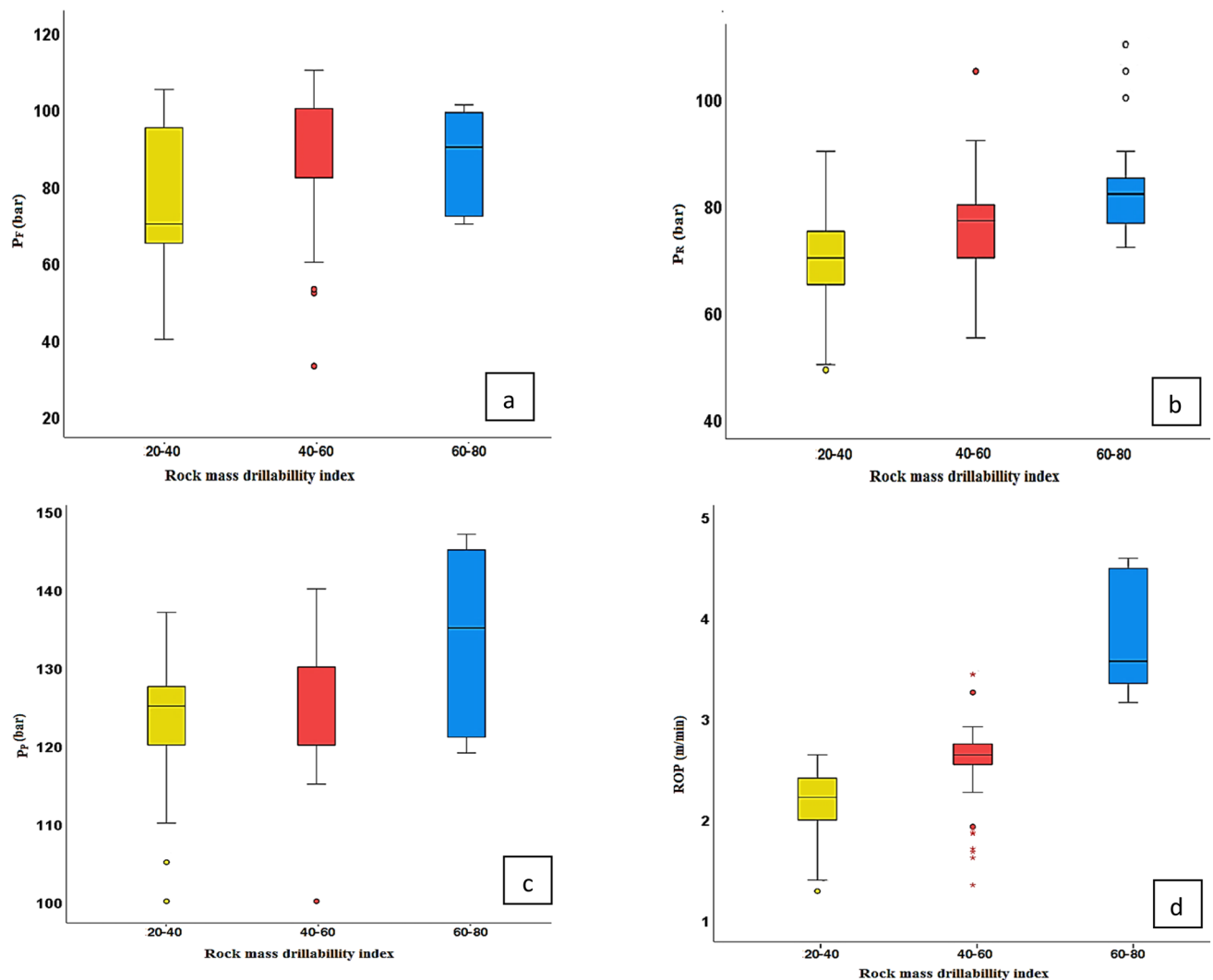
class was clustered. Each solid point in Fig. 3 represents the average of multiple blasting holes, ranging from 32 to 75, with one point representing each group.

### Data preprocessing

To ensure accurate model training, it is crucial to preprocess raw data due to its high level of noise and outlier data. Neglecting to remove outliers and reduce noise can impede the model learning process and prolong training time. At this stage, the collected data undergoes review and analysis to prepare for program entry, with 80% of the data used for training and 20% for testing randomly. To ensure accurate data analysis, it is critical to take two necessary actions: First, data analysis, verification of their accuracy and precision and Second, data matching; to avoid scattering and place all the data in a specific interval. Non-quantitative data can be handled using various techniques. Improper scaling can cause regression analysis to misestimate the significance of each variable, and deleting or averaging out-of-range data can be used as a solution. This study employs the IQR method, one of the most commonly used outlier labeling techniques. To handle extreme values that can behave as outliers in a noisy dataset, it is necessary to perform outlier labeling before processing. The IQR method is considered one of the best approaches, which takes the interquartile range (IQR) into account for labeling outliers, using Eq. (1):

$$\text{IQR} = Q_3 - Q_1 \quad (1)$$

In Eq. (1),  $Q_3$  represents the 3rd quartile, and  $Q_1$  represents the 1st quartile of the data. To determine the upper and lower limits of the extreme boundaries, the IQR value is multiplied by a factor and then subtracted from  $Q_1$  and added to  $Q_3$ . The most commonly used factor for this is 1.5. To analyze the sample data outliers, box diagrams were drawn according to the drilling parameter and rock mass drillability index classification that is shown in Fig. 4. In Fig. 4, under different Rock mass Drillability index, all four drilling parameters contained outliers. To



**Figure 4.** Drilling parameter box diagrams for each RDi. (a) Feed pressure ( $P_F$ , bar); (b) Rotation pressure ( $P_R$ , bar); (c) Hammer pressure ( $P_H$ , bar); (d) The rate of penetration (ROP, m/min).

eliminate the effects of the outliers, the average value of the parameter was used to replace the outlier of data in this paper. The changes in data characteristics before and after cleaning are compared in Table 4, after cleaning.

After data cleaning; to avoid scattering and place all the data in a specific interval, in the field of machine learning, different evaluation indexes (that is, different features in feature vectors are described as different evaluation indexes) often have different dimensional and dimensional units, which will affect the results of data analysis. In order to eliminate the dimensional influence between indexes, data normalization is required.

Normalization is a technique used to limit input data to a specific range, often between 0 and 1. In this study, the data were normalized using the min–max normalization method, which is defined in Eq. (2):

$$x' = (x - x_{\min}) / (x_{\max} - x_{\min}) \quad (2)$$

In Eq. (2),  $x$  represents the value of the original data,  $x_{\max}$  represents the maximum value of the original data,  $x_{\min}$  represents the minimum value of the original data, and  $x'$  represents the normalized value.

### Validation of the models

In computational mechanics, several metrics are available in statistics. This research implements five performance metrics. The testing performance is assessed based on five statistical performance criteria, namely determination coefficient ( $R^2$ ), variance accounted for (VAF), mean absolute error (MAE), mean absolute percentage error (MAPE), and root mean square error (RMSE). VAF can be used to evaluate the variance proportion of the variables, while RMSE, MAE, and MAPE are frequently used to compare the prediction errors of different models—the lower the RMSE, MAE, and MAPE, the better the model performs. These statistical indices are expressed as follows (Eqs. 3–6)<sup>71–73</sup>:

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^n (T_i - O_i)^2}{n}} \quad (3)$$

$$\text{MAE} = \frac{\sum_{i=1}^n |T_i - O_i|}{n} \quad (4)$$

$$\text{MAPE} = \frac{100}{n} \times \sum_{i=1}^n \left| \frac{T_i - O_i}{T_i} \right| \quad (5)$$

$$\text{VAF} = \left[ 1 - \frac{\text{var}(T_i - O_i)}{\text{var}(T_i)} \right] \times 100 \quad (6)$$

where var denotes the variance,  $T_i$  and  $O_i$  are the measured and predicted values,  $n$  is the sample size.

### A penetration rate model for Jumbo drills

#### Drilling performance assessment using non-linear regression analyses

Assessment of geo-mechanical properties of rock formations plays a crucial role in determining the efficacy of rock drilling, as well as predicting drilling-related costs, timing, and productivity. Predictive modeling techniques, such as regression analysis, can provide valuable insights into the relationships between independent and dependent variables.

In order to ensure that a model is capable of making accurate predictions on unseen data, it is necessary for it to perform well not only on the data that it was trained on but also on novel data. To accomplish this, the

RD <sub>i</sub>	Index	P <sub>F</sub>		P <sub>R</sub>		P <sub>P</sub>		ROP	
		Before	After	Before	After	Before	After	Before	After
20–40	mean	80.83	80.90	69.96	69.75	125.51	125.61	2.07	2.08
	std	16.02	15.66	10.00	9.49	7.09	6.81	0.40	0.40
	min	40	50	49	50	100	105	1.28	1.28
	max	105	105	105	90	140	140	2.95	2.95
40–60	mean	85.05	86.92	78.36	78.61	124.76	125.30	2.74	2.78
	std	16.99	14.29	12.93	11.22	7.42	6.35	0.45	0.37
	min	33	38	55	55	100	110	1.34	1.85
	max	110	110	115	102	145	145	3.84	3.69
60–80	mean	87.15	88.88	82.62	80.15	133.78	133.45	3.78	3.81
	std	11.69	10.88	9.15	4.42	11.01	10.73	0.52	0.51
	min	70	70	72	72	119	119	3.15	3.15
	max	101	101	110	90	147	147	4.58	4.58

**Table 4.** Comparison of the data characteristics before and after cleaning.

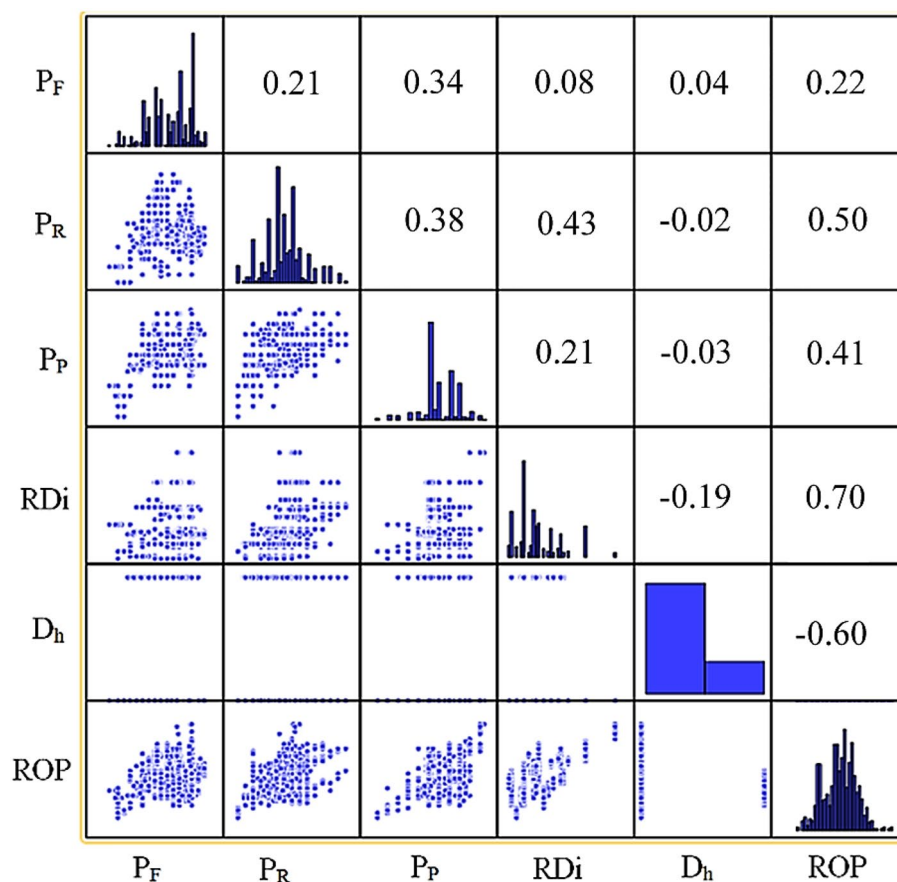
dataset is partitioned randomly into two subsets, with 80% of the data used for training and the remaining 20% used for testing.

Initially, we conducted several basic regression analyses utilizing linear, logarithmic, power, and exponential functions to establish the correlation between each independent variable and the rate of penetration. The coefficients of determination resulting from these basic regression analyses are compiled in Table 5. Based on the findings of the correlation analysis between each independent variable and ROP, it became evident that the rate of penetration cannot be anticipated by a single variable alone. Instead, it is impacted by numerous factors. Consequently, statistical methods and machine learning methods were used to attain more precise predictions of the rate of penetration. The correlogram of the dataset is represented in Fig. 5. In Fig. 5, scatter plots for each pair of variables are presented on the lower left, correlation coefficients are shown on the upper right, and histograms are placed in the middle. It is apparent from Fig. 5 that there was no substantial overlap between the independent variables, and none of them were removed while constructing the model.

In contrast to the limited scope of traditional linear regression, non-linear regression is capable of estimating models that encompass complex relationships between independent and dependent variables. In this paper, different algorithms were used to describe nonlinear multivariate regression functions, especially focusing on simple regression functions. The resulting models were then compared to identify the most accurate one based on the model development data. Table 6 presents five distinct non-linear multivariable regression equations that were derived using the model development data. As shown in Table 6, the correlation coefficients for these models ranged from 0.82 to 0.87 when assessed solely on the model development datasets. However, when considering the testing datasets, the correlation coefficients were found to be between 0.81 and 0.86.

Parameters	Linear	Logarithmic	Power	Exponential
$P_F$	0.13	0.19	0.25	0.18
$P_R$	0.45	0.49	0.53	0.46
$P_P$	0.27	0.28	0.34	0.31
$RD_i$	0.72	0.72	0.70	0.68

**Table 5.** Correlation analysis of rate of penetration and individual parameters.



**Figure 5.** Correlogram of the dataset.



Model	Regression equation	R <sup>2</sup> (Training)	R <sup>2</sup> (Testing)
1	$ROP = 2.39 \times \frac{P_F^{0.11} \times P_R^{0.3} \times P_P^{0.6} \times RD_i^{0.59}}{D_h^{1.73}}$	0.87	0.86
2	$ROP = 199P_F^{0.001} + 295P_R^{0.002} + 254P_P^{0.001} + 0.07RD_i^{0.006} - 744D_h^{0.005}$	0.85	0.82
3	$ROP = (0.003 P_F + 0.2)^2 + (0.007 P_R + 0.2)^2 + (0.008 P_P - 0.3)^2 + (0.02 RD_i + 0.006)^2 + (-0.07 D_h + 4.6)^2$	0.82	0.81
4	$ROP = 2.8 - \frac{19.6}{P_F} - \frac{49.7}{P_R} - \frac{192}{P_P} - \frac{73.8}{RD_i} + \frac{216.3}{D_h}$	0.85	0.84
5	$ROP = 2.824 \times 1.002^{P_F} \times 1.004^{P_R} \times 1.004^{P_P} \times 1.012^{RD_i} \times 0.97^{D_h}$	0.83	0.81

**Table 6.** Correlation coefficients for non-linear multivariable regression models.

By examining the developed nonlinear regression models, it was found that the model specified by the below equation is the most accurate model among the models.

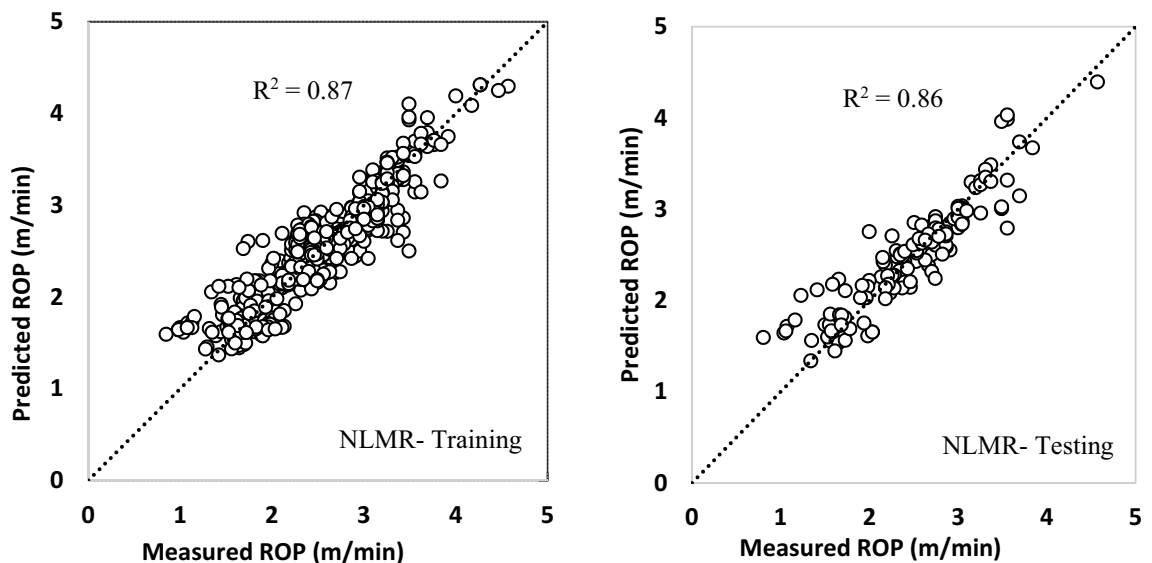
$$ROP = 2.39 \times \frac{P_F^{0.11} \times P_R^{0.3} \times P_P^{0.6} \times RD_i^{0.59}}{D_h^{1.73}} \tag{7}$$

The cross-correlation graph of the nonlinear multiple regression model is presented in Fig. 6.

### Drilling performance assessment using machine learning methods

Model complexity significantly influences the trade-off between bias and variance. Bias arises when complex relationships are oversimplified, and variance measures the sensitivity to dataset variations. In ROP modeling, an analytical model assuming a power law for drilling speed based on a few variables can introduce notable bias due to oversimplification. Analytical ROP models tend to exhibit high bias as they oversimplify the complex drilling process. In contrast, machine learning algorithms allow for the inclusion of numerous input variables. However, increasing model complexity raises the risk of overfitting, which adds variance. Machine learning algorithms provide flexibility in exploring drilling parameters by adapting to the data's characteristics via algorithm-specific hyperparameters rather than predefined equations<sup>4</sup>. Artificial neural networks, support vector machines, and random forests were chosen as the algorithms chosen for this paper based on performance and applicability to a wide range of problems.

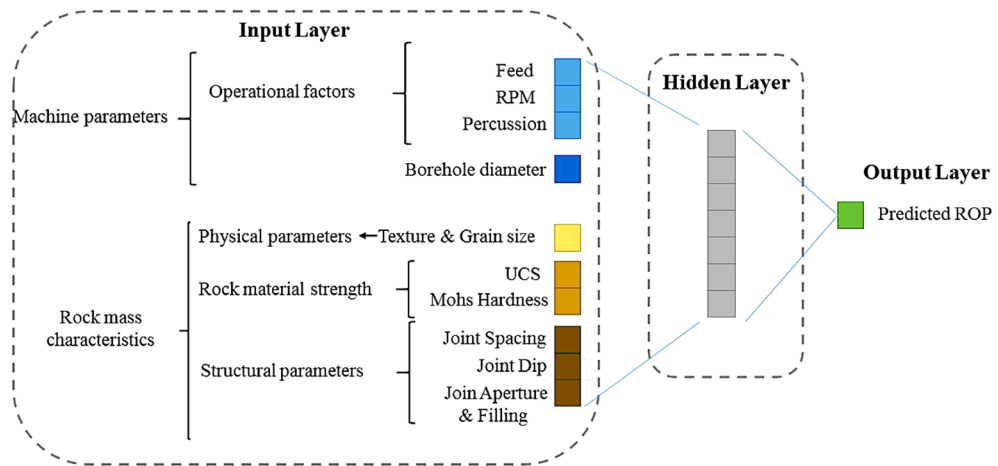
The architectures of machine learning models can be controlled by hyperparameters, but there are no hard and fast rules for selecting the optimal hyperparameters, as the ideal model structure can differ depending on the application. To find the best hyperparameter combinations, researchers often define a grid and employ cross-validation techniques. The current study also uses this methodology. To obtain an optimal network structure with the most suitable hyperparameters, multiple networks were constructed and their results were compared to determine the best one. The overall outcomes of the finest constructed networks are presented in Table 7. The selection and prediction of the ANN model involve several important factors including feature selection, network architecture, and transfer of functions across layers, along with the choice of the training algorithm. Eventually, a network with 7 hidden layers (as depicted in Fig. 7) was chosen. The transfer function used in the hidden layers is hyperbolic tangent, while the output layer employs the exponential transfer function. Cross-correlations graphs of MLP, SVR, and RF models are presented in Fig. 8. The results and performance indices of developed models are presented in Table 8.



**Figure 6.** The cross-correlation graph of the nonlinear multiple regression model.

Multilayer perceptron Grid		Best hyperparameter
Solver	SGD, Adam, LBFGS	LBFGS
Number of Neurons in the Hidden layer	3, 4, 5, 6, 7, 8, 9, 10, 11	7
Activation function	Logistic, Tanh, exponential	tanh
l2 Regularization ( $\alpha$ )	0.0001, 0.001, 0.01, 0.1	0.001
Support vector machines grid		Best Hyperparameter
Kernel function	Linear, 3rd Degree Polynomial, Gaussian	Gaussian
Epsilon ( $\epsilon$ )	0.01, 0.1, 1, 10	0.1
Budget (C)	1, 10, 100	10
Kernel coefficient ( $\gamma$ )	0.1, 0.2, 0.5	0.2
Random Forests Grid		Best Hyperparameter
Number of Trees	50, 100, 150, 300, 500	150
Number of Features	3, 4, 5	5

**Table 7.** Hyperparameter grid search for Multilayer perceptron, support vector machines and random forests ROP models.



**Figure 7.** The architecture of the MLP model used in the paper.

### Sensitivity analysis

To recognize the most sensitive factors affecting penetration rate, the cosine amplitude method (CAM) was utilized. To apply this method, all the data pairs are expressed in a common X-space which is used to construct a data array  $X$ , defined as<sup>74</sup>:

$$X = \{X_1, X_2, X_3, \dots, X_n\} \tag{8}$$

Each of the elements,  $X_i$ , in the data array  $X$  is a vector of lengths of  $n$ , that is:

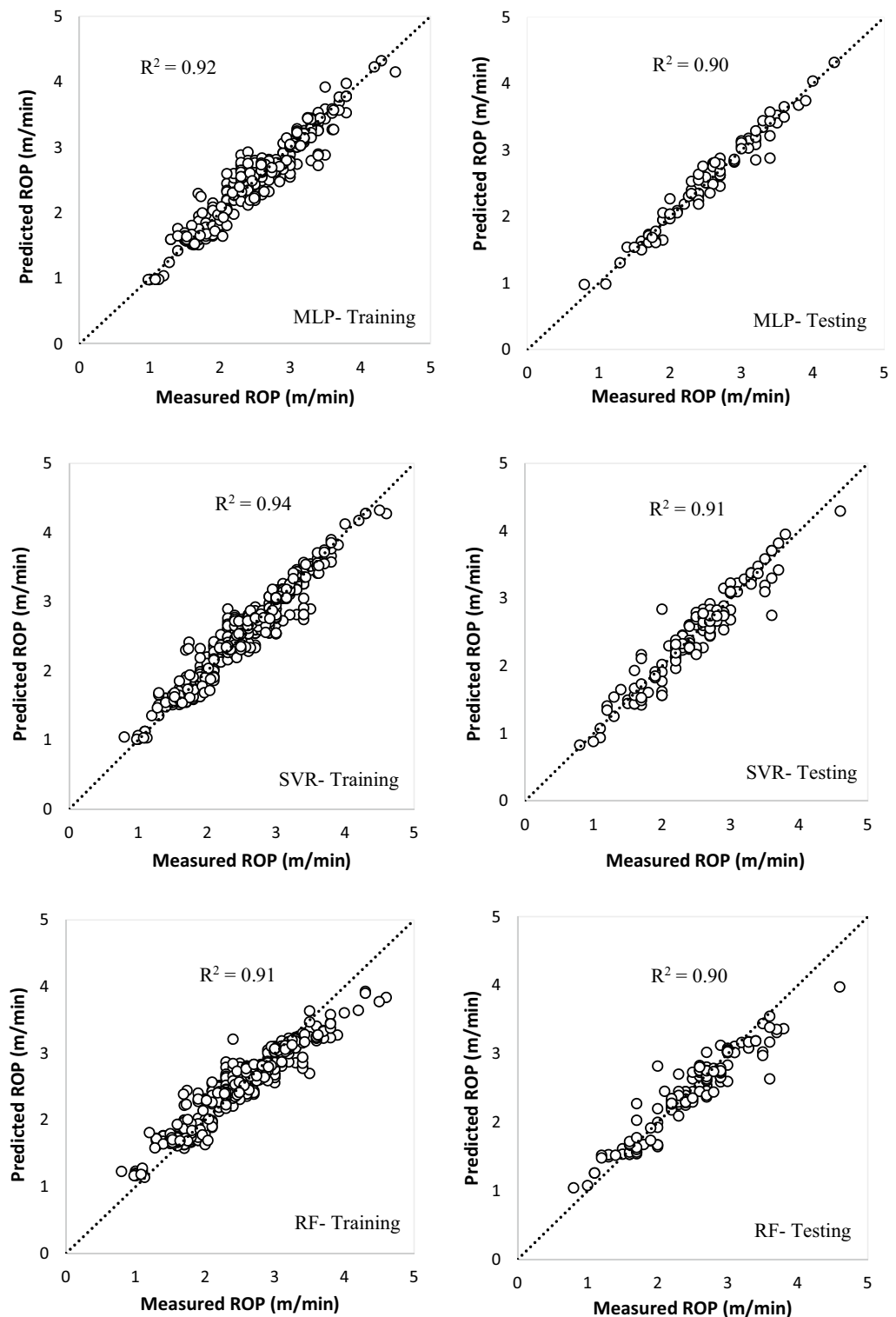
$$X_i = \{X_{i1}, X_{i2}, X_{i3}, \dots, X_{in}\} \tag{9}$$

Thus, each of the data pairs can be thought of as a point in  $n$  dimensional space, in which each point requires  $n$  coordinates for a full description. Each element of the relation,  $r_{ij}$ , results in a pairwise comparison of two data pairs. The strength of the relation between the data pairs,  $x_i$  and  $x_j$ , is given by the membership value expressing the strength:

$$R_{ij} = \frac{\sum_{k=1}^n (x_{ik} \times x_{jk})}{\sqrt{\sum_{k=1}^n x_{ik}^2 \times \sum_{k=1}^n x_{jk}^2}} \tag{10}$$

where  $i, j$ , and  $k$  represent respectively the counters of the number of input indicators in each data series, the indicators or factors related to each data series, and the number of data series or samples. The closer  $R_{ij}$  is to one, the greater the impact of the input indicator on the target index. If there is no effect, the value of  $R_{ij}$  will be zero.

The strengths of relations ( $r_{ij}$  values) between the ROP and input parameters are shown in Fig. 9. Considering that the values of  $R_{ij}$  are high for all parameters affecting the penetration rate, and based on the graph presented in Fig. 9, it can be concluded that all of the considered parameters are significantly involved in the rate of



**Figure 8.** Correlations of predicted ROP versus the measured ROP for machine learning models.

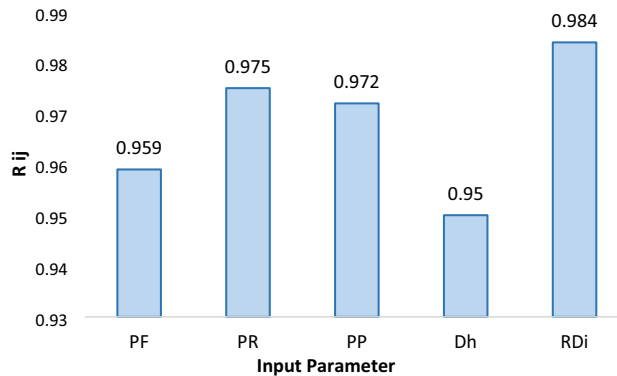
penetration. As it is shown, the most effective parameter on the ROP is Rock mass drillability index. In other words, the characteristics of the rock mass have a great influence on the drilling speed.

### Results and discussion

The application of an index that reflects the general characteristics of the rock mass can be instrumental in expediting the estimation of rock mass characteristics and reducing costs. The rock mass drillability index (RDi) possesses this ability and can prove beneficial in penetration rate studies. However, before utilizing any index

Model type	Training set					Testing set				
	R <sup>2</sup>	RMSE	MAE	MAPE	VAF	R <sup>2</sup>	RMSE	MAE	MAPE	VAF
NLMR	0.87	0.25	0.19	8.81	85.17	0.86	0.25	0.19	9.8	83.45
MLP	0.92	0.16	0.11	5.34	92.36	0.90	0.22	0.14	6.94	91.29
SVR	0.94	0.15	0.11	4.84	94.13	0.91	0.19	0.13	6.02	91.11
RF	0.91	0.20	0.15	6.60	90.52	0.90	0.22	0.16	7.20	89.49

**Table 8.** The results and performance indices of developed models.



**Figure 9.** Strengths of relation ( $R_{ij}$ ) between the penetration rate and each input parameter.

or parameter in the field of modeling for prediction, it is imperative to establish the effect of that parameter on the output result. Since the RD<sub>i</sub> has not been implemented in underground mines and underground drilling techniques differ from open pit mining, it is crucial to first demonstrate the efficacy of the RD<sub>i</sub> in underground drilling and investigate its effect on the penetration rate. After studying the relationship between the penetration rate and RD<sub>i</sub>, the index was considered as one of the inputs of the model that represents the characteristics of the rock mass.

The dataset utilized for calculations comprises 737 rows. Initially, missing values are eliminated. Subsequently, the IQR method is applied to detect outliers and bad data readings, and they are replaced with the parameter's average value to neutralize their effects. The resultant dataset comprises 737 rows, which are normalized using the min–max normalization method. Thereafter, the dataset is randomly split into 80–20% train and test subsets. The machine learning hyperparameters are fine-tuned utilizing the grid search range strategies, and the optimal hyperparameters are listed in Table 7.

Various metrics can be employed to evaluate the accuracy of the desired model in approximation. For instance, smaller values of MAE, MAPE, and RMSE indicate a higher accuracy in approximation, whereas bigger values of R<sup>2</sup> and VAF indicate the same. Furthermore, the MAPE values can be used to calculate the absolute value of the average percent relative error, which provides a comprehensive analysis of the model's performance. In addition, R<sup>2</sup> is instrumental in determining the percentage of the model outputs that can be defined by the fitted line on the data points, and a value of R<sup>2</sup> close to one is indicative of good accuracy in approximation.

The development of regression models and subsequent comparison between them revealed that the relationship between the penetration rate and its predictive parameters is nonlinear. Due to the potential for complex nonlinear relationships that regression models cannot extract between data, methods based on artificial intelligence were employed to develop the model. Upon analysis of the training and test data results, it was discovered that machine learning models exhibit higher accuracy and lesser error than regression models. This highlights the veracity of the assumption regarding the existence of complex relationships between the parameters that predict penetration rate. Thus, it can be inferred that mathematical models are comparatively less accurate than computational intelligence-based models. The mathematical models assume that the effect of some drilling variables on ROP has a linear and absolute incremental behavior. Upon scrutinizing the models generated through the machine learning method, it is evident that the SVR method outperforms the other models in terms of higher accuracy and lesser error.

To compare the results obtained from the machine learning (ML) models with the previous Rate of Penetration (ROP) models, a ROP model (Shen model) was selected, the mathematical relations of which are depicted in Eq. (11)<sup>68</sup>.

$$v = -0.77p_t(0.5p_c^2 - 15.7p_c + 117.6) \quad (11)$$

where  $v$  denotes the rate of penetration (mm/s),  $p_t$  is Propelling pressure (MPa) and  $p_c$  is Percussive pressure (MPa).

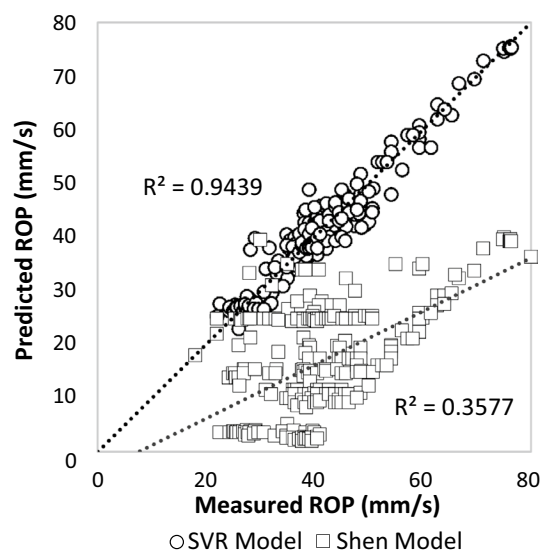
A comparison was then conducted between these models using the database. Further details regarding these models can be found in the relevant research literature<sup>68</sup>. Figure 10 illustrates the outcomes of the previous model, demonstrating the prediction error. It is evident from the figure that this model poorly fits the data, exhibiting high error and significant deviation in ROP values. This inadequacy stems from the model's failure to account for the effect of rock mass discontinuities. Also, the model developed in this research provides better results due to the use of more parameters. Based on the observations from Fig. 10, it can be inferred that the mathematical models are notably less accurate when contrasted with the computational intelligence-based models.

## Conclusions

Efficient drilling processes can lead to cost savings through increased penetration rates. Optimizing drilling processes requires a thorough understanding of the interplay between various parameters affecting the process. Estimating the penetration rate factor enables engineers to better plan for the future and adjust drilling parameters for optimal results. To prepare a suitable database, effective parameters affecting the Rate of Penetration (ROP), including rock mass properties and machine specifications for 737 boreholes in the mentioned underground mines, were collected and datasets were divided randomly into training (80%) and testing (20%) datasets. These parameters served as model inputs to predict ROP. By utilizing machine parameters and rock properties, multiple regression analyses and machine learning methods were implemented, revealing their dominant effect on jumbo drill performance. These algorithms attempt to learn the physical behavior between independent and dependent parameters based on the underlying theory. The learned relation can then be generalized to predict system behavior. In summary, the following outcomes can be drawn from this research:

- Several simple regression analyses were conducted, determining correlations between different variables and penetration rate. Results showed that determination coefficients of simple regression analyses were rather low.
- Regression models were developed with high correlation coefficients. The regression model for jumbo drill penetration rate (Eq. 7) can effectively predict rock drill performance by utilizing rock properties and machine parameters. Other proposed models (Table 6) for predicting penetration rate were also found to be reliable.
- Machine learning algorithms demonstrated better results than regression algorithms for predicting ROP.
- The Support Vector Regression (SVR) model exhibited good predictive results for both training and testing databases compared to other machine learning models. SVR also had the lowest error rate compared to Random Forest (RF) and Multi-Layer Perceptron (MLP) methods. However, the use of these methods may be conditional.
- According to the sensitivity analysis of the effective parameters in the penetration rate, it can be concluded that all of the considered parameters significantly contribute to the rate of penetration. The most influential parameter on the ROP is the Rock Mass characteristic.

The Rock Drillability Index (RDi) being unaffected by rock type suggests that the prediction models developed in this study can estimate jumbo drill performance across different challenging rock conditions. While our model shows promise for application in dry and semi-dry rock formations, it's essential to recognize its limitations in saturated rock conditions and high-water-volume mines, as these scenarios may introduce complexities not accounted for in the model, potentially affecting prediction accuracy. Additionally, the model's applicability may



**Figure 10.** Correlations of predicted ROP versus the measured ROP for Shen model vs. SVR model using data set.

be limited in less jointed rock masses, particularly those lacking cohesion, such as homogeneous rock formations, posing challenges for accurately predicting drilling rates. Furthermore, our model is tailored specifically for underground drilling conditions, and its effectiveness in scenarios involving drilling in open pits or larger diameter holes may be compromised due to distinct challenges not addressed in its design. The prediction models established in this research can be further refined, suggesting significant possibilities for drilling automation in underground mines. Hence, the findings of this study hold great potential for the field of drilling automation in underground mines.

## Data availability

The datasets used and analyzed during the current study available from the corresponding author on reasonable request.

Received: 5 January 2024; Accepted: 15 April 2024

Published online: 18 April 2024

## References

- Akbar, A. & Altindag, R. An investigation of different drill bits and drilling angles in blast hole drilling. *Rock Eng. Rock Mech.* <https://doi.org/10.1201/b16955-178> (2014).
- Darbor, M., Faramarzi, L. & Sharifzadeh, M. Performance assessment of rotary drilling using non-linear multiple regression analysis and multilayer perceptron neural network. *Bull. Eng. Geol. Environ.* **78**, 1501–1513. <https://doi.org/10.1007/s10064-017-1192-3> (2019).
- Najjarpour, M., Jalilfar, H. & Norouzi-Apourvari, S. Half a century experience in rate of penetration management: Application of machine learning methods and optimization algorithms: A review. *J. Pet. Sci. Eng.* **208**, 1–18. <https://doi.org/10.1016/j.petrol.2021.109575> (2022).
- Soares, C. & Gray, K. Real-time predictive capabilities of analytical and machine learning rate of penetration (ROP) models. *J. Petrol. Sci. Eng.* **172**, 934–959. <https://doi.org/10.1016/j.petrol.2018.08.083> (2019).
- Elmgerbi, A. M., Ettinger, C. P., Tekum, P. M., Thonhauser, G. & Nascimento, A. Application of machine learning techniques for real time rate of penetration optimization. In *Proceedings of the SPE/IADC Middle East Drilling Technology Conference and Exhibition, Abu Dhabi, United Arab Emirates 25–27*. <https://doi.org/10.2118/202184-MS> (2021).
- Cheniany, A., Khoshrou, S. H., Kourosh, S. & Khademi Hamidi, J. An estimation of the penetration rate of rotary drills using the Specific Rock Mass Drillability index. *Int. J. Min. Sci. Technol.* **22**, 187–193. <https://doi.org/10.1016/j.ijmst.2011.09.001> (2012).
- Singh, D. P. A study of some aspects of drag-bit drilling. *Met. Min. Rev.* **5**(15), 1–4 (1976).
- Pathankar, A. G. & Misra, G. B. Drillability of rocks in percussive drilling from 'energy per unit volume' as determined with a microbit. *Min. Eng.* **32**, 1407–1410 (1980).
- Ersoy, A. & Walker, M. D. Prediction of drill bit performance using multivariate linear regression analysis. *Trans. Inst. Min. Metall.* **104**, 70–94 (1995).
- Wilmot, G. M., Calhoun & Perrin, B. V. P. Formation drillability: Definition, quantification, and contribution to bit performance evaluation. In *SPE/IADC Middle East Drilling Technology Conference, Abu Dhabi, UAE, Paper SPE 57558*. <https://doi.org/10.2118/57558-MS> (1999).
- Yarali, O. & Kahraman, S. The drillability assessment of rocks using the different brittleness values. *Tunnel. Undergr. Space Technol.* **26**, 406–414. <https://doi.org/10.1016/j.tust.2010.11.013> (2011).
- Bilim, N. Determination of drillability of some natural stones and their association with rock properties. *Sci. Res. Essays* **6**, 382–387. <https://doi.org/10.5897/SRE10.878> (2011).
- Selim, A. A. & Bruce, W. E. Prediction of penetration rate for percussive drilling. USBM. RI 7396 (1970).
- Clark, G. B. Principles of rock drilling. *Colo School Mines Q* **74**, 91–93 (1979).
- Howarth, D. F. & Rowland, J. C. Quantitative assessment of rock texture and correlation with drillability and strength properties. *Rock Mech. Rock Eng.* **20**, 57–85. <https://doi.org/10.1007/BF01019511> (1987).
- Thuro, K. & Spaun, G. Introducing the destruction work as a new rock property of toughness referring to drillability in conventional drill- and blast tunnelling. In *ISRM International Symposium* (1996).
- Kahraman, S. Rotary and percussive drilling prediction using regression analysis. *Int. J. Rock Mech. Min. Sci.* **36**, 981–989. [https://doi.org/10.1016/S0148-9062\(99\)00050-9](https://doi.org/10.1016/S0148-9062(99)00050-9) (1999).
- Taheiri, A., Qao, Q. & Chanda, E. Drilling penetration rate estimation using rock drillability characterization index. *J. Inst. Eng. Ser. D* **97**, 159–170. <https://doi.org/10.1007/s40033-015-0104-6> (2016).
- Paone, J., Madson, D. & Bruce, W. E. Drillability studies - laboratory percussive drilling. US Bureau of Mines RI 7300 (1969).
- Tandanand, S. & Unger, H. F. Drillability determination: A drillability index of percussive drills. US Bureau of Mines RI 8073 (1975).
- Pathankar, A. G. & Misra, G. B. A critical appraisal of the Protodyakonov Index. *Int. J. Rock Mech. Min. Sci. Geomech.* **13**(8), 249–251. [https://doi.org/10.1016/0148-9062\(76\)91545-X](https://doi.org/10.1016/0148-9062(76)91545-X) (1976).
- Rabia, H. *Oilwell Drilling Engineering: Principles and Practise* (Graham & Tromtman, 1985).
- Bond, D. F. The optimization of PDC bit selection using sonic velocity profiles present in the Timor Sea. *SPEDE* **5**, 135–142 (1990).
- Perrin, V. P. Wilmot, M. G. & Alexander, W. L. Drilling Index: A new approach to bit performance evaluation. In *SPE/IADC Drilling Conference, Paper SPE 37595* 199–205 (1977).
- Norwegian Institute of Technology. *Drillability, Drilling Rate Index Catalogue* 8–79 (Norwegian Institute of Technology, 1981).
- Schmidt, R. L. Drillability studies: Percussive drilling in the field. USBM RI 7684 (1972).
- Hoseinie, S. H., Aghababaei, H. & Pourrahimian, Y. Development of a new classification system for assessing of rock mass drillability index (RDi). *Int. J. Rock Mech. Min. Sci.* **45**(1), 1–10. <https://doi.org/10.1016/j.ijrmms.2007.04.001> (2008).
- Hoseinie, S. H., Ataei, M. & Osanloo, M. A new classification system for evaluating rock penetrability. *Int. J. Rock Mech. Min. Sci.* **46**, 1329–1340. <https://doi.org/10.1016/j.ijrmms.2009.07.002> (2009).
- Hoseinie, S. H., Ghorbani, S. & Ghodrati, B. Selection of suitable drilling method in Razgah nepheline syenite mine, a systematic approach. *Equip. Mater.* **1**, 56–60. <https://doi.org/10.17580/em.2020.01.11> (2020).
- Eskikaya, S., Bilgin, N. & Dincer, T. A model to predict the cutting performance of rapid excavation systems. In *Proceedings of the Seventh International Symposium on Mine Plan Equipment Selection* (ed. Singhal, R. K.) 575–9 (1998).
- Cheniany, A., Hasan, K. S., Shahriar, K. & Khademi Hamidi, J. An estimation of the penetration rate of rotary drills using the specific rock mass drillability index. *Int. J. Min. Sci. Technol.* **22**(2), 187–193. <https://doi.org/10.1016/j.ijmst.2011.09.001> (2012).
- Hustrulid, W. A. & Fairhurst, C. A theoretical and experimental study of the percussive drilling of rock. Part I: The SPEY of percussive drilling. *Int. J. Rock Mech. Min. Sci. Geomech. Abstracts* **8**, 35–56. [https://doi.org/10.1016/0148-9062\(71\)90045-3](https://doi.org/10.1016/0148-9062(71)90045-3) (1971).
- Thuro, K. Prediction of drillability in hard rock tunnelling by drilling and blasting. In *Proc ITA Congress Tunnels for People* (eds Golser, J. et al.) 103–108 (Balkem, 1997).



34. Fish, B. G. The basic variables in rotary drilling. *Mine Quarry Eng.* **27**, 74–81 (1968).
35. Bauer, A. & Calder, P.N. Open pit drilling—factors influencing drilling rates. In *Proceedings of the 4th Canadian Symposium on Rock Mechanics* 1–33 (1967).
36. Lebel, J. R. *An Investigation to Evaluate the Relationship Between Rock Quality Index (RQI) and Powder Factor for Surface Mining* Dissertation (University of British Columbia, 1984).
37. Bauer, A. & Crosby, W. A. Drilling application. In *Section 612 Surface Mining* 2nd edn (ed. Kennedy, B. A.) 524–539 (Society for Mining Metallurgy and Exploration Inc Port City Press, 1990).
38. Workman, L. & Szumanski, M. Which blasthole rig? In *Proceedings Surface Mining* (ed. Glen, H. W.) (SAIMM, 1996).
39. Workman, L. & Szumanski, M. Which blasthole rig?. *World mining Equipment* 23–27 (1997).
40. Rabia, H. Specific energy as a criterion for drill performance prediction. *Int. J. Rock Mech. Min. Sci. Geomech. Abstracts* **19**, 39–42. [https://doi.org/10.1016/0148-9062\(82\)90709-4](https://doi.org/10.1016/0148-9062(82)90709-4) (1982).
41. Clark, G. B. Principles of rock drilling part-1. *Colo School Mines* **77**, 118 (1982).
42. Bilgin, N., Eskikaya, S. & Dincer, T. The performance analysis of large diameter blast hole rotary drilling in Turkish Coal Enterprises. In *Mine mechanization and automation* (eds Almgren, G. et al.) (Balkema, 1993).
43. Kahraman, S., Balci, C., Yazici, S. & Bilgin, N. Prediction of the penetration rate of rotary blast hole drills using a new drillability index. *Int. J. Rock Mech. Min. Sci.* **37**(5), 729–743. [https://doi.org/10.1016/S1365-1609\(00\)00007-1](https://doi.org/10.1016/S1365-1609(00)00007-1) (2000).
44. Kahraman, S., Bilgin, N. & Feridunoglu, C. Dominant rock properties affecting the penetration rate of percussive drills. *Int. J. Rock Mech. Min. Sci.* **40**, 711–723. [https://doi.org/10.1016/S1365-1609\(03\)00063-7](https://doi.org/10.1016/S1365-1609(03)00063-7) (2003).
45. Kahraman, S., Ipek, M., Guleryuz, U. & Gunaydin, N. Performance prediction of a jumbo drill in Pozanti-Ankara Motorway Tunnel (Turkey). *Tunnel. Undergr. Space Technol.* **21**, 265–265. <https://doi.org/10.1016/j.tust.2005.12.126> (2006).
46. Altindag, R. Evaluation of drill cuttings in prediction of penetration rate by using coarseness index and mean particle size in percussive drilling. *Geotechn. Geol. Eng.* **22**, 417–425. <https://doi.org/10.1023/B:GEGE.0000025043.92979.48> (2004).
47. Singh, T. N., Gupta, A. R. & Sain, R. A comparative analysis of cognitive system for the prediction of drillability of rocks and wear factor. *Geotech. Geol. Eng.* **24**, 299–312. <https://doi.org/10.1007/s10706-004-7547-0> (2006).
48. Ataei, M., Kakaei, R., Ghavidel, M. & Saeidi, O. Drilling rate prediction of an open pit mine using the rock mass drillability index. *Int. J. Rock Mech. Min. Sci.* **73**, 130–138. <https://doi.org/10.1016/j.ijrmms.2014.08.006> (2015).
49. Hung, N. V., Gerbaud, L., Souchal, R., Urbanczyk, C. & Fouchard, C. Penetration rate prediction for percussive drilling with rotary in very hard rock. *J. Sci. Technol.* <https://doi.org/10.15625/0866-708X/54/1/5956> (2016).
50. Navarro, J. *The Use of Measurme While Drilling for Rock Mass Characterization and Damage Assessment in Blasting*, Doctoral Thesis (Madrid University, 2016).
51. Zhao, S., Wang, M., Yi, W., Yang, D. & Tong, J. Intelligent classification of surrounding rock of tunnel based on 10 machine learning algorithms. *Appl. Sci.* **12**, 1–20. <https://doi.org/10.3390/app12052656> (2022).
52. Koopialipoor, M., Tootoonchi, H. & Jahed Armaghani, D. Application of deep neural networks in predicting the penetration rate of tunnel boring machines. *Bull. Eng. Geol. Environ.* **78**, 6347–6360. <https://doi.org/10.1007/s10064-019-01538-7> (2019).
53. Grima, M. A. & Babuska, R. Fuzzy model for the prediction of unconfined compressive strength of rock samples. *Int. J. Rock Mech. Min. Sci.* **36**, 339–349. [https://doi.org/10.1016/S0148-9062\(99\)00007-8](https://doi.org/10.1016/S0148-9062(99)00007-8) (1999).
54. Yan, C. et al. Predicting TBM penetration rate with the coupled model of partial least squares regression and deep neural network. *Rock Soil Mech.* <https://doi.org/10.16285/j.rsm.2020.5164> (2021).
55. Bermudez, S. A. *Integer Programming to Evaluate Operational Impact of Penetration Rate Predictive Model* Doctoral thesis (McGill University, 2020).
56. Yang, J. F., Zhai, Y. J., Xu, D. P. & Han, P. SMO algorithm applied in time series model building and forecast. *Proc Sixth Int Conf Mach Learn Cybern ICMLC* **4**, 2395–2400. <https://doi.org/10.1109/ICMLC.2007.4370546> (2007).
57. James Pacis, F. et al. Improving predictive models for rate of penetration in real drilling operations through transfer learning. *J. Comput. Sci.* <https://doi.org/10.1016/j.jocs.2023.102100> (2023).
58. Xuan-dong, H., Shi, X., Wu, Y. & Tang, L. Application of neural network model with partial least square regression in prediction of peak velocity of blasting vibration. *J. Vib. Shock* **32**(12), 45–49 (2013).
59. Kidega, R., Nelima Ondiaka, M., Maina, D., Arap Too Jonah, K. & Kamran, M. Decision based uncertainty model to predict rockburst in underground engineering structures using gradient boosting algorithms. *Geomech. Eng.* **30**(3), 259–272. <https://doi.org/10.12989/gae.2022.30.3.259> (2022).
60. Muhammad Shahani, N., Kamran, M., Zheng, X. & Liu, C. Predictive modeling of drilling rate index using machine learning approaches: LSTM, simple RNN, and RFA. *Pet. Sci. Technol.* **40**(5), 534–555. <https://doi.org/10.1080/10916466.2021.2003386> (2022).
61. Kamran, M., Muhammad Shahani, N. & Jahed Armaghani, D. Decision support system for underground coal pillar stability using unsupervised and supervised machine learning approaches. *Geomech. Eng.* **30**(2), 107–121. <https://doi.org/10.12989/gae.2022.30.2.107> (2022).
62. Arehart, R. Drill-bit diagnosis with neural networks. *SPE Comput. Appl.* **2**, 24–28. <https://doi.org/10.2118/19558-PA> (1990).
63. Bilgesu, H., Tetrick, L., Altimis, U., Mohaghegh, S. & Ameri, S. A new approach for the prediction of rate of penetration (ROP) values. In *SPE Eastern Regional Meeting* (Society of Petroleum Engineers, 1997). <https://doi.org/10.2118/39231-MS>.
64. Hegde, C., Hugh, D., Millwater, H. & Gray, K. Analysis of rate of penetration (ROP) prediction in drilling using physics-based and data-driven models. *J. Pet. Sci. Eng.* **159**, 295–306. <https://doi.org/10.1016/j.petrol.2017.09.020> (2017).
65. Barbosa, L. F., Nascimento, A., Hugo Mathias, M. & Carvalho, J. A. Machine learning methods applied to drilling rate of penetration prediction and optimization: A review. *J. Pet. Sci. Eng.* **183**, 1–20. <https://doi.org/10.1016/j.petrol.2019.106332> (2019).
66. Kamran, M. A Probabilistic approach for prediction of drilling rate index using ensemble learning technique. *J. Min. Environ.* **12**, 327–337. <https://doi.org/10.22044/jme.2021.10689.2030> (2021).
67. Muhammad Shahani, N., Kamran, M. & Zheng, X. Predictive modeling of drilling rate index using machine learning approaches: LSTM, simple RNN, and RFA. *Pet. Sci. Technol.* **40**, 534–555. <https://doi.org/10.1080/10916466.2021.2003386> (2022).
68. Shen, Q., Wang, Y., Cao, R. & Liu, Y. Efficiency evaluation of a percussive drill rig using rate-energy ratio based on rock drilling tests. *J. Pet. Sci. Eng.* <https://doi.org/10.1016/j.petrol.2022.110873> (2022).
69. Su, O. Performance evaluation of button bits in coal measure rocks by using multiple regression analyses. *Rock Mech. Rock Eng.* **49**, 541–553. <https://doi.org/10.1007/s00603-015-0749-3> (2016).
70. Siva Prasad, B. N. V., Murthy, V. S. M. R. & Naik, S. Compensious Index for Drillability: A rapid tool to assess drill penetration rate and bit life for rock engineering applications. *Bull. Eng. Geol. Environ.* **82**, 79. <https://doi.org/10.1007/s10064-023-03081-y> (2023).
71. Hosseini, S., Pourmirzaee, R., Jahed Armaghani, D. & Sabri Sabri, M. M. Prediction of ground vibration due to mine blasting in a surface lead–zinc mine using machine learning ensemble techniques. *Sci. Rep.* **13**, 6591. <https://doi.org/10.1038/s41598-023-33796-7> (2023).
72. Hosseini, S. et al. Assessment of the ground vibration during blasting in mining projects using different computational approaches. *Sci. Rep.* **13**, 18582. <https://doi.org/10.1038/s41598-023-46064-5> (2023).
73. Khatti, J. & Grover, K. S. CBR Prediction of pavement materials in unsoaked condition using LSSVM, LSTM-RNN, and ANN approaches. *Int. J. Pavement Res. Technol.* <https://doi.org/10.1007/s42947-022-00268-6> (2023).

74. Yang, Y. & Zhang, Q. A hierarchical analysis for rock engineering using artificial neural networks. *Rock Mech. Rock Eng.* **30**, 207–222. <https://doi.org/10.1007/BF01045717> (1997).

### Author contributions

S.H: Data curation, Writing—original draft.S.H H: Conceptualization, Methodology.R.B: Investigation, Supervision.

### Competing interests

The authors declare no competing interests.

### Additional information

**Correspondence** and requests for materials should be addressed to S.H.H.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2024