

## An Intelligent Approach to Predicting Dilution, Overbreak and Costs in Underground Mining Using Kolmogorov-Arnold Networks



Jhon Kener Vera-Encarnación<sup>\*ID</sup>, Marco Antonio Cotrina-Teatino<sup>ID</sup>, Solio Marino Arango-Retamozo<sup>ID</sup>

Department of Mining Engineering, Faculty of Engineering, National University of Trujillo, Trujillo 13001, Peru

Corresponding Author Email: [jkverae@unitru.edu.pe](mailto:jkverae@unitru.edu.pe)

Copyright: ©2025 The authors. This article is published by IETA and is licensed under the CC BY 4.0 license (<http://creativecommons.org/licenses/by/4.0/>).

<https://doi.org/10.18280/mmep.120308>

### ABSTRACT

**Received:** 3 January 2025

**Revised:** 5 March 2025

**Accepted:** 10 March 2025

**Available online:** 31 March 2025

#### **Keywords:**

*costs, dilution, prediction, KAN, overbreak*

The main objective of this research was to predict dilution, overbreak, and costs in underground mining using Kolmogorov-Arnold Networks (KAN). The KAN model was rigorously evaluated against other machine learning approaches, including Random Forest (RF), Extreme Gradient Boosting (XGBoost), and Multilayer Perceptron (MLP). The methodology included the collection of 732 records, which was used for training and testing the models using the k-fold technique. The results indicated that the KAN model exhibited the highest performance, particularly in dilution prediction, with a Mean Squared Error (MSE) of 0.001, a Root Mean Squared Error (RMSE) of 0.032, a Mean Absolute Error (MAE) of 0.055, and an  $R^2$  of 0.999. Regarding overbreak, KAN achieved an MSE of 0.002, an RMSE of 0.045, an MAE of 0.004, and an  $R^2$  of 0.984. For cost, it achieved an MSE of 0.001, an RMSE of 0.032, a MAE of 0.027, and an  $R^2$  of 0.999, reaffirming its position as the most efficient model. Additionally, the integration of a multi-objective particle swarm optimization algorithm enabled the optimization of drilling and blasting parameters, resulting in notable reductions: 11.3% in operational dilution, 25% in overbreak, and 7.6% in operational costs. This study demonstrates the effectiveness of KANs in predicting key mining indicators, improving efficiency, safety, and profitability.

## 1. INTRODUCTION

Drill and blast continue to stand out as one of the most versatile and widely employed techniques for excavation in mining, quarrying, and tunneling in hard rock, thanks to its relatively low cost, efficiency, and ease of implementation [1-3].

However, this method has some disadvantages, such as the destruction of the surrounding rock mass, which can result in unwanted cavities due to blasting, referred to as “overbreak” [4]. Furthermore, in mining, it is confronted with dilution, defined as the percentage of material or waste with economic value removed beyond the delimited area of extraction, whether planned or unplanned. Unplanned dilution not only reduces the quality of the extracted ore but also increases processing and transportation costs, which negatively impacts profits, underlining the need for effective management [5, 6].

Mine production capacity and profitability are directly dependent on ore loss and dilution indicators, which are key factors in assessing the development prospects of a mining company. Therefore, assessing the geological and economic factors associated with ore dilution and loss is critical at the feasibility stage of mine design, as well as during mining operations [7]. The rate of ore dilution varies based on the mining method employed for underground extraction of mineral resources [8-10].

Ensuring integrity and quality in mineral resource utilization is the key driver of annual productivity growth in

mining enterprises and the expansion of industrial reserves. However, factors such as mine structural complexity, ore and surrounding rock instability, and adverse hydrogeological conditions negatively impact mining productivity. These factors can lead to an increase in ore dilution of up to 35%-40% and ore loss of up to 25%, highlighting the need for mitigation measures [11].

The problem arises in an underground mine, where traditional approaches are severely constrained, as evidenced by the large differences between predicted values and those observed in the field. Some critical factors behind this problem include errors in drill and blast design, such as incorrect drill hole placement, inappropriate application of explosives, and excessive charge length. In general, empirical estimates tend to be insufficient for properly modeling rock mass behavior because they do not include detailed analyses, such as numerical modeling or in-depth in-situ testing, which negatively affects operational planning and overall profitability.

Such discrepancies lead to uncontrolled and excessive fragmentation of the rock mass, causing overbreak, which leads to dilution and unwanted extraction of waste material, increasing the operational costs associated with clearing, support, loading, and hauling.

Therefore, predicting and mitigating overbreak and dilution during the mining process are considerable challenges for any mining project. To improve these predictions, there have been attempts to incorporate more advanced approaches that

employ artificial intelligence along with more sophisticated mathematical modeling techniques [12, 13].

Recent studies have indicated substantial improvements in accuracy. Jorquera et al. [14] used various machine learning (ML) models to estimate dilution, achieving a precision score of 0.835, an accuracy of 0.804, and an area under the curve of 0.942 with the Random Forest model.

On the other hand, Chimunhu et al. [15] employed geological, geotechnical, and mine design variables available in the early stages of sub level open stopping to predict dilution. They achieved 93% accuracy in dilution prediction using the principal component analysis-classification and regression trees (PCA-CART) model.

Zhao and Niu [16] used a backpropagation neural network to estimate unplanned mineral dilution, overcoming the limitations of the empirical ELOS graph method and achieving a determination coefficient ( $R^2$ ) greater than 0.95. Jang et al. [17] developed a neuro-fuzzy system to predict unplanned dilution and mineral loss, obtaining a correlation coefficient (R) of 0.72.

Models ranging from simple to hybrid approaches have been used for overbreak prediction, with the latter showing superior performance. Jang and Topal [18] utilized 49 datasets of Rock Mass Rating (RMR) and overbreak as input and output variables, respectively, to predict overbreak, achieving an  $R^2$  of 0.945 using an Artificial Neural Network (ANN).

To forecast overbreak and manage it prior to drilling and blasting activities, they implemented an ANN alongside a hybrid system combining a genetic algorithm and an artificial bee colony algorithm with an ANN, yielding an  $R^2$  exceeding 0.90 [19-22]. Furthermore, models incorporating fuzzy logic and adaptive neuro-fuzzy inference systems integrated with particle swarm optimization have been developed, delivering predictions with an  $R^2$  above 0.94 [23, 24].

Due to the non-linear and multifactorial nature of mining operations such as blasting, multi-objective intelligent optimization algorithms have been implemented to address these challenges. Bakhtavar et al. [25] mainly reduced fragmentation size and costs through a multi-objective stochastic planning model. The multi-objective particle swarm optimization method (MOPSO) has evolved as a highly effective technique for optimizing multiple indices in blasting operations, where several objectives can be improved simultaneously [26, 27].

This study uses Kolmogorov-Arnold networks to predict dilution, overbreak, and costs in underground mining. It combines geological, geomechanical, and operational data in a deep learning model to improve economic viability. This new approach integrates theory with deep learning, enhancing efficiency and optimization for sustainable mining.

This research follows the following structure: Section 2 presents the materials and methodology, including the fundamentals of the KAN model. Section 3 discusses the model's performance and results. Finally, Section 4 provides conclusions and recommendations for further research.

## 2. MATERIALS AND METHODS

### 2.1 Data collection and analysis

The database was from an underground mine in northern Peru, which consisted of 732 records. These records included geometric, geological, geomechanically and operational

parameters, as well as the costs associated with each unit operation. This stage was critical, as the most important variables were identified in order to train the model and predict the three targets of the project.

The variables for the creation of the predictive model were: mining width (MW), working height (WH), vein thickness (VT), dip, RQD, RMR, charge length (CL), stemming (S), advance length (AL), explosive quantity (EQ), broken volume (BV), tonnage (TN), charge factor (CF), power factor (PF), operating dilution (OD), overbreak (OB), drill and blast cost (CBD), clean-up cost (CC), loading cost (LC), haulage cost (HC), support cost (SC) and total operating cost (TOC).

Table 1 presents a statistical analysis of 732 observations per variable. The main characteristics of the operational and geomechanically parameters considered in the model are presented.

**Table 1.** Input and target statistics

Variables	Mean	Var.	Std.	Min.	Max.
MW (m)	2.39	0.00	0.03	2.34	2.44
WH (m)	2.19	0.00	0.03	2.14	2.24
VT (m)	0.45	0.01	0.09	0.30	0.60
Dip	47.54	2.17	1.47	45.00	50.00
RQD	65.05	12.50	3.54	60.03	69.97
RMR	55.06	49.49	7.04	45.00	65.00
CL (m)	1.24	0.00	0.01	1.22	1.26
S (m)	0.38	0.00	0.01	0.36	0.40
AL (m)	1.46	0.00	0.03	1.41	1.51
BV (m <sup>3</sup> )	6.18	0.01	0.11	5.97	6.37
TN (Tn)	16.37	0.08	0.29	15.88	16.88
EQ (kg)	15.73	0.32	0.57	14.71	16.71
CF (kg/m <sup>3</sup> )	2.55	0.01	0.10	2.31	2.77
PF (kg/Tn)	0.96	0.00	0.04	0.88	1.05
OD (%)	81.23	13.61	3.69	74.55	87.67
OB (m <sup>3</sup> )	0.29	0.01	0.11	0.08	0.48
CBD (US\$/Tn)	47.56	1.30	1.14	45.51	49.50
CC (US\$/Tn)	21.85	0.34	0.58	20.85	22.85
HC (US\$/Tn)	14.63	0.35	0.59	13.65	15.65
LC (US\$/Tn)	15.41	0.34	0.59	14.40	16.40
SC (US\$/Tn)	5.59	0.08	0.28	5.10	6.10
TOC (US\$/Tn)	105.03	2.49	1.58	100.93	109.53

The MW has an average of 2.39 m with a minimum variability ( $\pm 0.03$  m), as well as the height of the workings, which has an average of 2.19 m. The actual VT shows a larger dispersion with a mean of 0.45 m and a standard deviation of  $\pm 0.09$  m. The dip reaches a mean value of 47.54°, with a range varying between 45° and 50°, reflecting its relative stability. As for the geomechanically parameters, the RQD and RMR register averages of 65.05 and 55.06, respectively, highlighting that the RMR has a greater dispersion ( $\pm 7.04$ ), which could significantly influence the mine planning. The target total operating cost averages 105.03 US\$/Tn, with main components being drilling and blasting (47.56 US\$/Tn), cleaning (21.85 US\$/Tn), hauling (14.63 US\$/Tn), loading (15.41 US\$/Tn) and support (5.59 US\$/Tn). These values show that drilling and blasting accounts for the largest percentage of the total cost. The targets, operational dilution, averages 81.23%, and overbreak, averages 0.29 m<sup>3</sup>, evidencing the need to optimize operations to reduce costs and minimize losses associated with unwanted material or excess fragmentation.

The detection of multicollinearity among the predictors was conducted through Variance Inflation Factor (VIF) analysis, given that excessive multicollinearity can affect model

stability and the interpretation of coefficients. A threshold of  $VIF \leq 10$  was established, where Work Area (WA) exhibited a value of 24.03 and Hydraulic Radius (HR) reached 22.07. Consequently, both were removed due to their high correlation with other variables, posing a risk of redundancy and result distortion.

The remaining variables had VIF values  $\leq 10$ , ensuring that they did not compromise model stability. Among them, CL (4.43), S (4.34), EQ (4.20), PF (3.12), and CF (3.08) showed moderate correlations but were deemed acceptable.

TN (3.06) and BV (2.05) were also retained in the analysis. The rest of the variables presented values  $< 2$ , including HC (1.04), WH (1.03), and RQD (1.02), indicating low collinearity. This selection ensures the model's robustness by reducing variance overestimation and increasing the reliability of the results.

Figure 1 illustrates the correlation matrix for the factors examined in the study. The primary factors correlated with OD are  $EQ=0.44$ ,  $PF=0.43$ , and  $CL=0.28$ . Furthermore,  $RMR=-0.55$  demonstrates a moderate negative correlation.

The data suggest that operational dilution is affected by a combination of blast design parameters (EQ, PF, and CL), the geomechanically properties of the RMR, and the results of the blasting process. OB exhibits the highest correlation with  $EQ=0.51$  and  $PF=0.52$ .

This analysis substantiates the idea that overbreak is influenced by both blast design parameters (EQ, PF) and the geomechanically properties of the  $RMR=-0.46$ . The variables that most significantly influence TOC are  $OB=0.67$ ,  $OD=0.64$ ,  $LC=0.39$ ,  $SC=0.27$ , and  $CDB=0.46$ .

These results highlight the importance of expenditures related to essential operating tasks, stressing the need to improve these elements to reduce overall costs in mining

operations.

The range of distribution of the data for each of the characteristics varied considerably. For example, RMR ranges from 45 to 65. OD ranges from 74.55 to 87.67. If feature scaling is not considered, the marked differences in the ranges could lead the machine learning algorithm to incorrectly assess the relative importance of each feature [28]. Therefore, before training the machine learning models, feature scaling was performed in order to normalize the distribution of the data in each feature. In this study, the normalization process was adopted, in which the values were adjusted and scaled to lie within the range of 0 to 1. Eq. (1) was used to find the normalization.

$$X' = \frac{X - X_{min}}{X_{max} - X_{min}} \quad (1)$$

The normalized value is represented by  $X'$ , the real value is  $X$ . The minimum and maximum values of the distribution for the selected features are defined by  $X_{min}$  and  $X_{max}$ . After data normalization, k-fold was applied. K-fold cross-validation is a fundamental method in machine learning and statistical modeling for assessing model performance. It is obtained by partitioning the data into k equally sized subsets, training the model on k-1 folds, and validating it on the remaining one, repeating the cycle k times [29]. This improves data utilization by minimizing bias and enhancing evaluation stability [30]. For our study, k-fold with  $k=5$  was used across all models, ensuring a balance between computational efficiency and validation robustness. Once the KAN model was implemented, it was compared with the three most commonly used ML models to predict the established objectives.

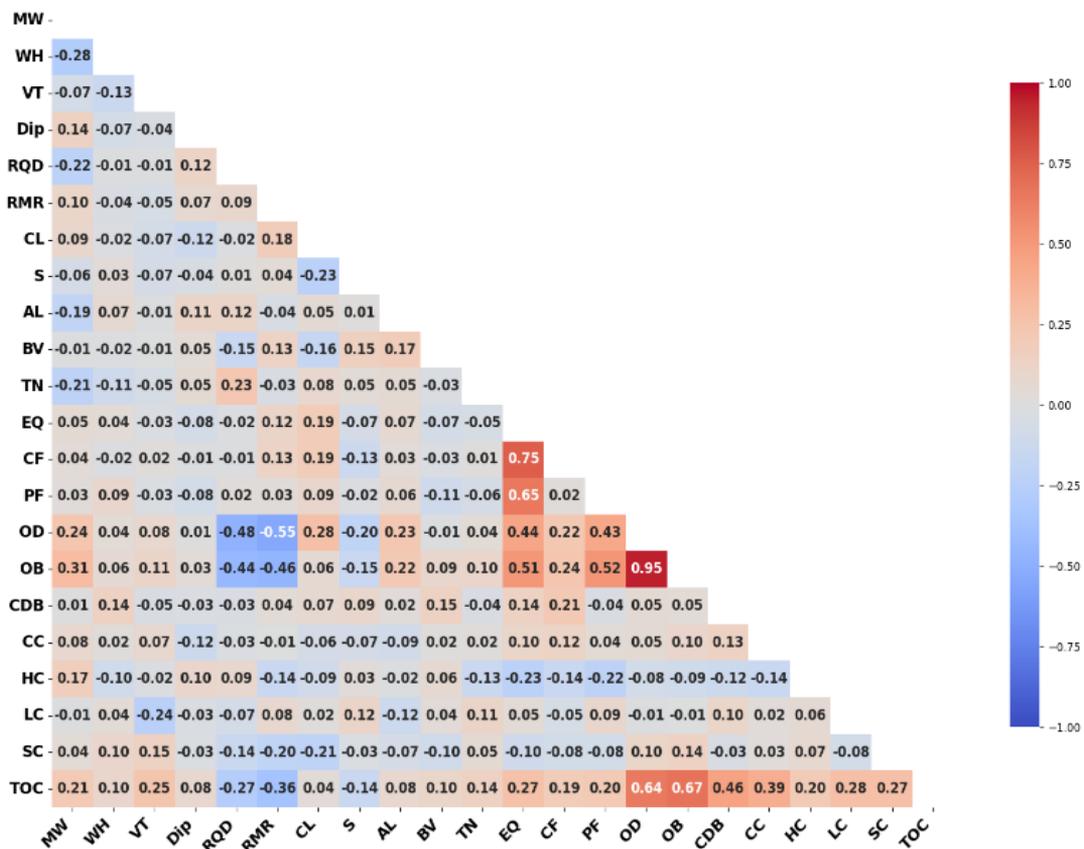


Figure 1. Correlation matrix of study variables

## 2.2 KAN

It is an innovative neural network architecture based on the Kolmogorov and Arnold theorems, which state that any continuous function of multiple variables can be expressed as the sum of single-variable functions [31].

Vladimir Arnold and Andrey Kolmogorov proved that any continuous multivariate function  $F$  defined on a bounded domain can be expressed as a finite combination of continuous univariate functions, linked through the binary operation of addition. This representation is particularly significant in the context of smooth operations, as it ensures that the involved functions are differentiable and possess mathematical properties that optimize their analysis and applicability in modelling complex systems across various scientific and engineering disciplines.

$$f(x) = f(x_1, \dots, x_n) = \sum_{q=1}^{2n+1} \Phi_q \left( \sum_{p=1}^n \Phi_{q,p}(x_p) \right) \quad (2)$$

Eq. (2) shows an external and an internal sum. The outer sum computes  $\sum_{q=1}^{2n+1} \Phi_q$ , from  $2n+1$  in terms of  $\Phi_q = \mathbb{R} \rightarrow \mathbb{R}$ . The inner sum computes  $n$  terms for each  $q$ , where  $\Phi_{q,p} = [0,1]^n \rightarrow \mathbb{R}$ . In this way, KAN networks are able to use single-variable functions that can be learned and incorporated instead of the fixed weights typical of traditional neural networks. This feature allows for better modelling and fitting of the network to complex patterns in the data.

All the weight parameters of this models are treated as functions of a variable and more often than not are parameterized by splines. Spline functions are piecewise polynomials and are defined by a group of control points which shape the boundary. These control points can be altered without restriction which enables the local contour of the function to be modified while its global behavior remains unaffected. This is the reason why spline functions, while being able to fit the complex patterns in the data, are able to retain the smoothness of the activation functions [32].

Not only does this enhance the flexibility of the network but also enhances its capability to capture and model highly nonlinear relationships [33]. This model has been subject to extensive validation and benchmarking and it is known to deliver solid performance in different domains, which proves the model efficiency and performance [31-39].

The KAN model based on B-Splines enables flexible and differentiable approximations of the nonlinear relationships implicit in the data. The KAN architecture is defined by a set of  $k$  nodes and a spline degree  $d$ , which determine the complexity of interpolation for each layer [31]. Each input is transformed through a combination of B-Spline basis functions, which are weighted by trainable parameters [40]. These basis functions divide the input domain into smooth and continuous regions, avoiding the saturation issues associated with traditional activation functions. The input values are evaluated in these basis functions during forward propagation, generating nonlinear representations that can be modified through training.

The KAN model for the project consists of a structure composed of two B-Spline layers with 512 and 256 nodes, respectively, followed by a dense layer with three output neurons, optimized to predict the target variables. The B-Spline layers are defined by a number of divisions ( $G$ ) determined by the knots and a polynomial degree ( $k$ ). The

optimization of the splines is performed iteratively using the Adam algorithm and by minimizing the MSE, adjusting the weights associated with the basic functions until the model achieves its best predictive capability. The established TensorFlow implementation ensures that the parameter update process remains efficient and stable during training. Below, a detailed pseudocode is provided, as shown in Figure 2, specifying the model structure and the calculation of B-Spline activations, ensuring that this pseudocode offers a clear representation of the training process.

```
# Define the B-Spline activation layer
DEFINE BSplineLayer(degree, num_knots, units):
    INITIALIZE knots in a fixed range [-1,1] (non-trainable)
    INITIALIZE weights randomly (trainable)

# Recursive B-Spline basis function
FUNCTION basis_function(x, i, d):
    IF d == 0:
        RETURN 1 if knots[i] ≤ x < knots[i+1], ELSE 0
    ELSE:
        COMPUTE term1 and term2 using recursive basis_function
        RETURN term1 + term2

# Forward pass of the layer
FUNCTION forward_pass(inputs):
    FOR each feature in inputs:
        COMPUTE combinations of basis functions
    RETURN output scaled by weights

# Define KAN model with B-Splines
DEFINE create_kan_model(input_shape, degree, num_knots):
    DEFINE inputs with input_shape
    APPLY BSplineLayer (512 units)
    APPLY BSplineLayer (256 units)
    DEFINE output layer (3 units, no activation)
    COMPILE model with Adam optimizer and MSE loss
    RETURN model

# Create and display model summary
kan_model = create_kan_model()
SHOW model summary
```

Figure 2. Pseudocode of the model

## 2.3 MLP

This type of neural network is created with an architecture that has an input layer, one or several hidden layers, and an output layer. Unlike a single-layer network such as the perceptron, multi-layer networks can learn and model complex non-linear relationships, enabling them to process data with complex and difficult patterns much more effectively [41].

$$\hat{y}(x) = \sum_{j=1}^H v_j f(w_j^T x + w_{bj}) + v_b \quad (3)$$

where,  $\hat{y}(x)$  represents the neural network's output corresponding to an input " $x$ ".  $H$  denotes the number of neurons in the hidden layer, while " $v_j$ " corresponds to the weights linking the hidden layer neurons to the output node. Each neuron in the hidden layer uses an activation function, referred to as " $f$ ". The term " $w_j^T x + w_{bj}$ " describes the weighted combination of the input " $x$ " where " $w_j$ ", are the weights and " $w_{bj}$ " is the bias term. The bias at the output of the network is represented by " $v_b$ ".

ReLU is the most common activation function used on this model because it is also the most common activation function used on deep neural networks these days given its simplicity to use [42]. In addition, the Adam optimization algorithm is also used, which is famous for its ability to adjust and optimize parameters while maintaining the stability of the system

during optimization [43].

## 2.4 RF

RF is a machine learning method that consists of many decision trees to improve the prediction score and robustness. Each tree is built around a certain portion of the training data set, which can be created using bootstrapping or random selects. Then the individual votes from the trees are added together, using mean value in regression works or majority vote in classification, to produce a more accurate and robust result [44]. Such measure decreases the model's variance and improves the model's ability to learn from complex noisy data [45].

In regression, the predictions are created by average the response variable of training instances from the leaf nodes. In classification, the output is voted on and whatever gets the majority wins. It is called aggregation with R. In regressions, the predictions from all trees are aggregated to form a single result but depending on the needed accuracy, "T" confirms set out level of accuracy.

$$\hat{y}(x) = \frac{1}{T} \sum_{t=1}^T \hat{y}^{(t)}(x) \quad (4)$$

## 2.5 XGBoost

This model utilizes the decision tree boosting methodology, renowned for its efficacy and robustness in addressing both regression and classification challenges. Its increasing popularity arises from its exceptional performance in data science competitions and its capacity to efficiently process massive datasets and handle complicated features. XGBoost is highly valued for its swift processing, robust generalization abilities, and effectiveness in optimizing models in high-dimensional or noisy datasets [46].

XGBoost improves decision tree efficacy through the application of boosting techniques. Unlike classic ensemble methods that train models independently, boosting combines models sequentially. Here, each new model is designed to address the errors of its predecessors, resulting in an ensemble that is both more precise and resilient [47].

Eq. (5) describes the prediction of  $\hat{y}_i$ , for an input  $x_i$  as the sum of  $K$  base functions  $f_k(x_i)$ .

$$\hat{y}_i = \Phi(x_i) = \sum_{k=1}^K f_k(x_i), f_k \in F \quad (5)$$

where, each  $f_k$  belongs to the functional space  $F$ , as defined in Eq. (6). The  $q(x)$  is a function that assigns each input "x" to a specific partition of the feature space, represented by "T". The expression  $w_{q(x)}$  represents the parameters associated with the partition defined by  $q(x)$ . Furthermore,  $m$  is the dimension of the input space, and  $t$  corresponds to the size of the parameter vector  $w$ .

$$F = \{f(x) = w_{q(x)}\}(q: R^m \rightarrow T, w \in R^t) \quad (6)$$

In summary, the equation models the prediction as a combination of parameterized base functions, where the structure of "F" defines how the parameters  $w$  and the input space partitions  $w_{q(x)}$  interacts to generate the predictions.

## 2.6 Performance metrics

The performance of each prediction model was assessed using four key metrics, defined as follows. MSE measures the average squared deviation between observed and predicted values. RMSE scales the errors to match the units of the actual data, enhancing interpretability. MAE computes the average absolute deviation between predicted and actual values, offering a straightforward measure of error magnitude, independent of direction. Values closer to 0 indicate a substantial improvement in the predictive performance of the machine learning methods [48]. On the other hand, the coefficient of determination ( $R^2$ ) ranges from 0 to 1, with higher values reflecting the model's ability to better explain the variability in the observed data [49].

$$MSE = \frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2 \quad (7)$$

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2} \quad (8)$$

$$MAE = \frac{1}{N} \sum_{i=1}^N |y_i - \hat{y}_i| \quad (9)$$

$$R^2 = 1 - \frac{\sum_{i=1}^N (y_i - \hat{y}_i)^2}{\sum_{i=1}^N (y_i - \bar{y}_i)^2} \quad (10)$$

where,  $y_i$  represents the observed value,  $\hat{y}_i$  corresponds to the predicted value,  $\bar{y}_i$  signifies the mean of the observed data, and  $N$  denotes the total number of samples in the training or testing phase.

## 2.7 Definition of hyperparameters

In this analysis, four machine learning models were used, each with different architectures and optimization approaches. KAN is based on B-Splines and is designed to approximate nonlinear relationships in a flexible and differentiable manner. Its design consists of an input layer with 19 nodes, two hidden layers with 512 and 256 neurons, respectively, and an output layer with three neurons. This model employs the B-Spline activation function ( $k=3, G=12$ ), is optimized using the Adam algorithm, and utilizes a batch size of 32.

Meanwhile, MLP consists of three hidden layers, each with 512 neurons. This model uses the ReLU activation function and is trained for 300 epochs with a batch size of 32, also using the Adam optimizer. Additionally, the RF algorithm was implemented, which is based on constructing an ensemble of 300 decision trees. This model is configured with a minimum of 10 samples per node and a maximum depth of 10, along with a minimum of 5 samples per leaf, providing good stability and generalization capability.

Finally, the XGBoost method was applied, which is based on a boosting approach over trees. This model optimizes performance by performing 500 iterations and has a learning rate of 0.1. Its maximum depth is set to 6, and it is configured with a random state of 42 to ensure result reproducibility (refer to Table 2).

**Table 2.** Hyperparameters of the models

Model	Hyperparameters	Value
KAN	Input layers	19
	Hidden layers	2
	Neurons per hidden layer	512-256
	Epochs	300
	Output layer	3
	Optimizer	Adam
	Batch size	32
	Activation function	B-Spline, k=3, G=12
	Input layer	19
	Hidden layers	3
MLP	Neurons in hidden layer	512-512-512
	Optimization	Adam
	Output layer	3
	Activation function	ReLU
	Training epochs	300
RF	Batch size	32
	Number of trees	300
	Random seed	42
	Minimum samples to split	10
	Maximum depth	10
XGBoost	Minimum samples per leaf	5
	Number of iterations	500
	Learning rate	0.1
	Maximum depth	6
	Random state	42

**3. RESULTS AND DISCUSSION**

A total of 732 records were processed, with 19 critical variables selected for the prediction of dilution, overbreak, and operational costs. The models were evaluated on different datasets using k-fold, and their results are detailed below.

**3.1 Training, validation and test by model**

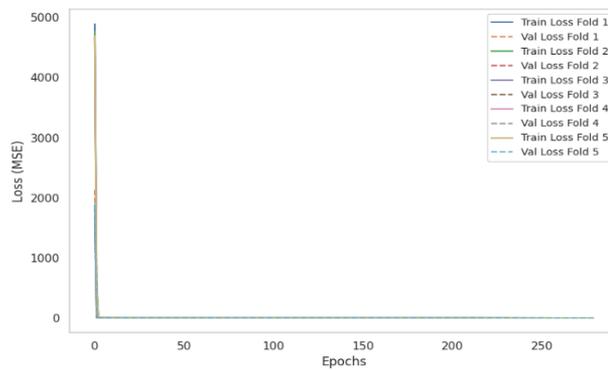
Figure 3 shows the loss function process during training and validation over 300 epochs for the KAN model. The training and validation lines are observed for each fold. Both lines decrease rapidly and then stabilize, converging towards constant values. This behavior indicates that the model is not experiencing overfitting, as the training and validation losses remain similar, suggesting an adequate generalization capability and accuracy on unseen data.

Figure 4 presents three scatter plots comparing the KAN model's predictions with actual values across the k-fold phases. The results show high correlations, suggesting that the model is robust and generalizes effectively. For dilution, overbreak, and cost, the correlation is nearly perfect (0.99) across all phases, with data points closely following the ideal regression line, indicating precise learning. Furthermore, the results for each fold remain consistent without significant variance, reinforcing the conclusion that the model is reliable and does not exhibit overfitting, demonstrating its ability to generalize to new data.

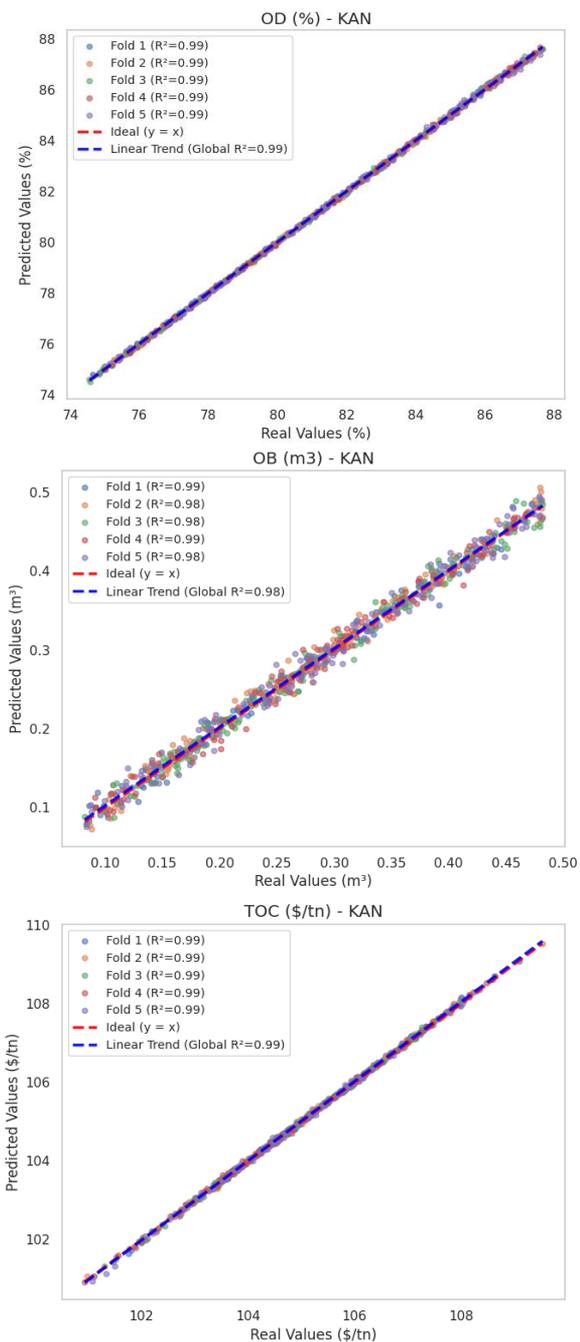
Figure 5 presents the results of the XGBoost model. The results show high accuracy for each fold, with an overall R<sup>2</sup> of 0.97 for operational dilution, while for overbreak, the value is 0.99. For total operating cost, it demonstrates high accuracy with an overall R<sup>2</sup> of 0.94.

Figure 6 shows the results of the RF model. The overall R<sup>2</sup> is 0.91 for dilution, demonstrating its good performance in predicting this target. For overbreak, the model showed a slight decrease in performance for each fold, obtaining an R<sup>2</sup>

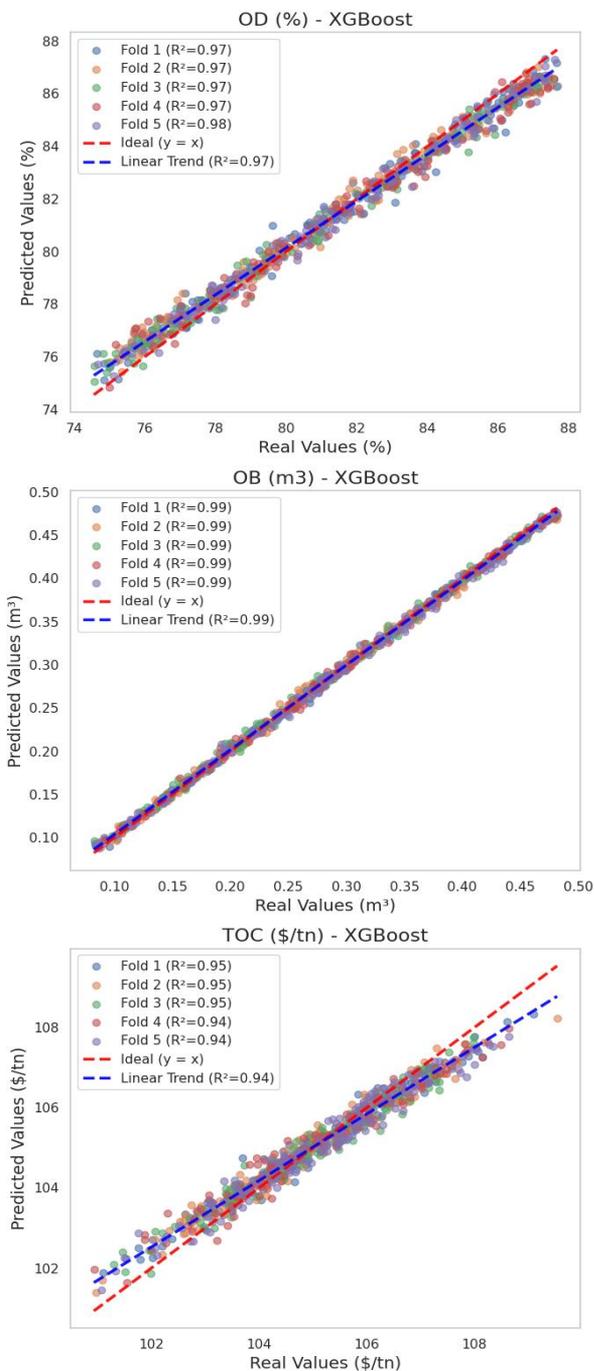
of 0.94. Similarly, in cost prediction, its performance declined for each fold, resulting in an overall R<sup>2</sup> of 0.84.



**Figure 3.** Training and validation curve – KAN



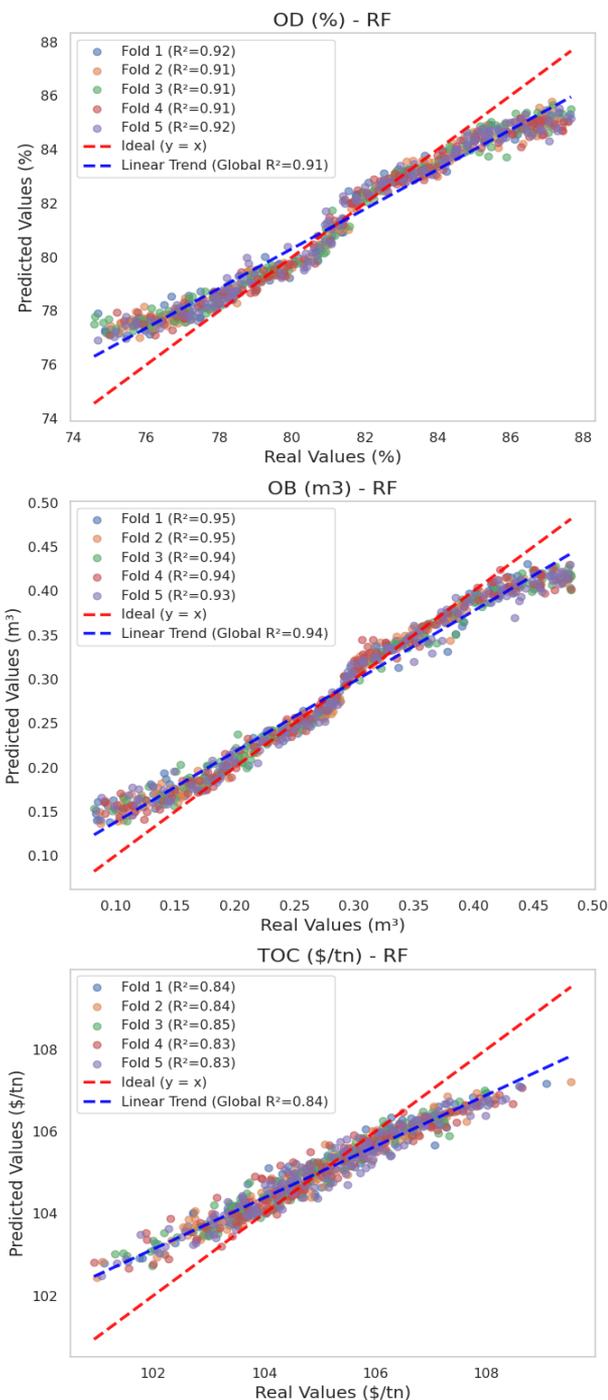
**Figure 4.** Evaluation of real vs. predicted values – KAN



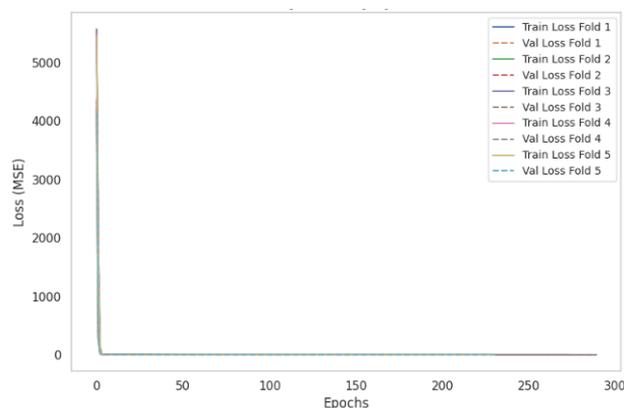
**Figure 5.** Evaluation of real vs. predicted values – XGBoost

Figure 7 depicts the evolution of the loss function during the training and validation phases of the MLP model, using MSE. The loss function shows a rapid initial decrease, followed by stabilization, as the training and validation curves align closely. This behavior suggests the model is free from overfitting, since the training loss does not significantly exceed the validation loss.

In Figure 8, the MLP model demonstrates strong performance across all variables. For operational dilution, the  $R^2$  is 0.99 for all datasets, indicating excellent correlation during training and high accuracy. For overbreak, it achieved an overall  $R^2$  of 0.93, reflecting solid overall performance with slight dispersion across folds, with the lowest values in folds 1 and 5, where  $R^2$  is 0.91. Finally, for total operating cost, the  $R^2$  values remain at 0.99 across all datasets, demonstrating robust performance.



**Figure 6.** Evaluation of real vs. predicted values – RF



**Figure 7.** Training and validation curve – MLP

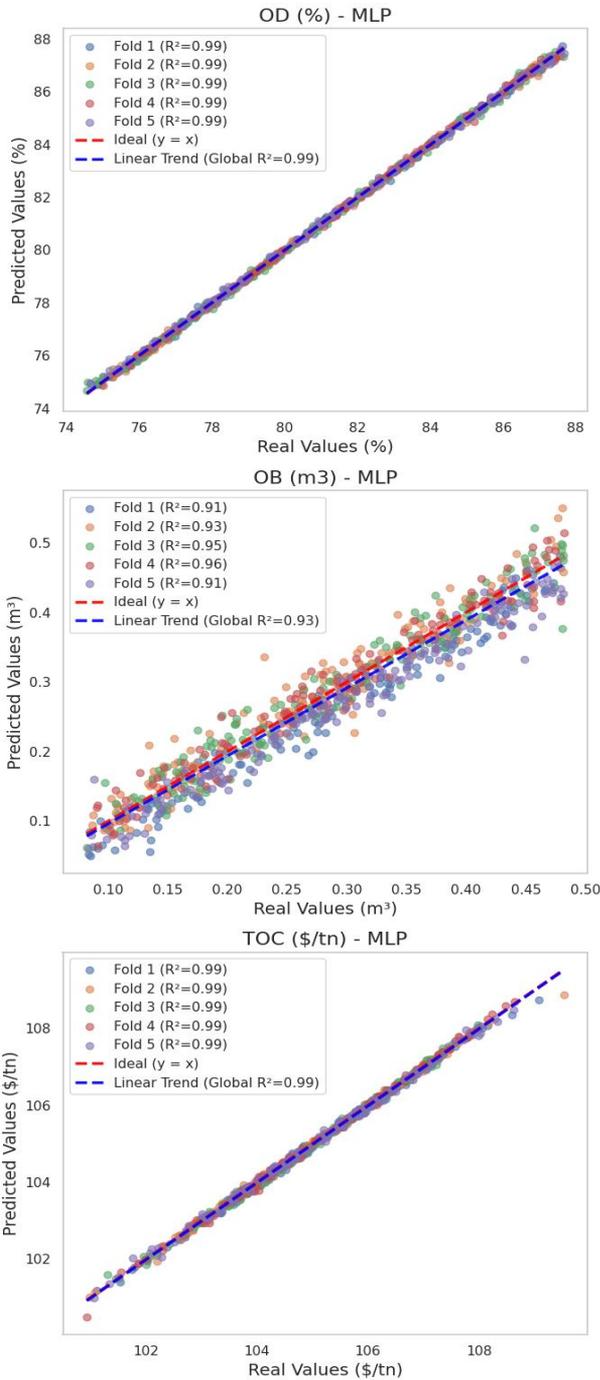


Figure 8. Evaluation of real vs. predicted values – MLP

### 3.2 Comparison and evaluation of the KAN model

Table 3 presents the analysis of the metrics. It is evident that the KAN model excels in all evaluations. In terms of dilution, KAN achieves the best values with MSE, RMSE, MAE, and  $R^2$  of 0.001, 0.032, 0.055, and 0.999, respectively, indicating an exceptional fit. For the overbreak metric, KAN has an MSE of 0.002 and RMSE of 0.045, an MAE of 0.004, and an  $R^2$  of 0.984. Regarding costs, KAN remains the leader with MSE, RMSE, MAE, and  $R^2$  of 0.001, 0.032, 0.027, and 0.999, respectively, reflecting the best fit among all models.

To rigorously evaluate the effectiveness of the implemented KAN model, both non-parametric and parametric statistical tests were conducted based on the data distribution. The Wilcoxon signed-rank test (WSRT) was applied to OD and OB, as these data do not follow a normal distribution. On the

other hand, for the TOC, which exhibits a normal distribution, the student's t-test was used.

Table 3. Performance metrics by model

Target	Metrics	KAN	MLP	XGBoost	RF
Operational dilution	MSE	0.001	0.019	0.285	1.177
	RMSE	0.032	0.138	0.534	1.085
	MAE	0.055	0.109	0.426	0.830
Overbreak	$R^2$	0.999	0.998	0.974	0.912
	MSE	0.002	0.114	0.001	0.132
	RMSE	0.045	0.333	0.032	0.363
Costs	MAE	0.004	0.242	0.003	0.219
	$R^2$	0.984	0.934	0.992	0.943
	MSE	0.001	0.013	0.132	0.405
	RMSE	0.032	0.114	0.363	0.636
	MAE	0.027	0.062	0.287	0.497
	$R^2$	0.999	0.995	0.943	0.837

Table 4. p-value result

Target	Comparison	p-value	Significance
Operational dilution	KAN vs. MLP	$3.22 \times 10^{-4}$	Yes
	KAN vs. RF	$5.23 \times 10^{-7}$	Yes
	KAN vs. XGBoost	$4.31 \times 10^{-4}$	Yes
Overbreak	KAN vs. MLP	$1.19 \times 10^{-6}$	Yes
	KAN vs. RF	$4.49 \times 10^{-6}$	Yes
	KAN vs. XGBoost	$8.32 \times 10^{-4}$	No
Costs	KAN vs. MLP	$4.12 \times 10^{-4}$	Yes
	KAN vs. RF	$4.78 \times 10^{-9}$	Yes
	KAN vs. XGBoost	$3.51 \times 10^{-4}$	Yes

The results of the WSRT test and student's t-test, presented in Table 4, reveal p-values (all below 0.05, except for XGBoost in overbreak), statistically validating significant distinctions between the KAN model and the alternative models evaluated. Regarding operational dilution, overbreak, and total operating cost, the KAN model demonstrated superior performance compared to MLP and RF. Conversely, XGBoost achieved better performance in predicting overbreak; however, KAN showed a significant difference in predicting dilution and cost. Therefore, the KAN model was chosen as the best-performing approach, outperforming MLP, XGBoost, and RF in the evaluated tasks.

As discussed previously, the KAN model outperformed other techniques based on various key metrics of performance. The superiority of the KAN model over other architectures lies in its use of learnable univariate functions. Unlike MLPs, which apply fixed activation functions such as ReLU or Tanh, KAN dynamically optimizes its activation functions through parameterized functions, such as splines, allowing for a more precise capture of complex relationships within the data. This leads to significantly more accurate models. Liu et al. [31] state that the primary strength of KAN architectures is their ability to increase the expressiveness of the approximation function by eliminating traditional fixed weights and inferring activation functions from univariate functions. This facilitates a better function fitting and helps solve partial differential equations with greater structural flexibility compared to MLP architectures [31, 32]. KAN also has the important ability to modify only the relevant regions without damaging prior knowledge and is less vulnerable to catastrophic forgetting during continuous learning; this latter feature is crucial in problems where stability and flexibility are essential. On the other hand, Ta [37] proposed the BSRBF-KAN architecture, which combines KAN, B-Splines, and radial basis functions

(RBF); this architecture can achieve better fitting in image classification problems, converging more stably and rapidly than MLPs and other KAN architectures, optimizing flexibility and interpretability. Moreover, in time series prediction problems, KAN models have outperformed MLPs and Recurrent Neural Networks (RNNs), achieving greater interpretability, efficiency, and accuracy while using significantly fewer parameters [34]. This establishes KAN as a remarkable architecture for modeling complex problems. In one study, Spline-KAN was used and achieved the same accuracy as MLPs but in half the training rounds, which translated into an improvement, despite a slight transition in computational time, as it demonstrated faster and more efficient convergence [38]. Similarly, Shukla et al. [39] analyzed the effectiveness of KAN and MLP in solving partial differential equations trained with noisy data; the results showed that when KAN was properly adjusted, it outperformed MLP in terms of accuracy and stability, enhancing KAN's capability in environments where flexibility, robustness, and computational efficiency are essential.

In this study, KAN achieved similar and superior performance compared to MLP, XGBoost, and RF models, while also providing advantages in both stability and flexibility in data representation. In this case, although MLP and XGBoost achieved high accuracy, KAN was chosen for its adaptability and ability to reach high levels of accuracy in the tasks performed.

### 3.3 Sensitivity analysis

The Sobol analysis is a sensitivity technique based on the concept of variance decomposition, which quantifies the effect that each input variable has on the model's response.

This procedure was used in the present research to characterize the effect of certain geomechanically and operational parameters on two key variables in underground mining: dilution behavior and overbreak behavior. The variables used for the analysis were defined along with their respective ranges, and samples were generated using the Saltelli method (5000 samples), which allows for a robust estimation of sensitivity indices. The analysis was also complemented by the use of normalized values in the range [0,1], which express the relative influence of each variable within the response of the model.

The first-order Sobol indices ( $S_1$ ) were calculated to determine the effect of each input variable on the output, while the second-order indices ( $S_2$ ) were computed to characterize interaction effects between pairs of input variables on the model's response. The total-order indices (ST) were calculated, including interactions with other input variables, even those corresponding to first- and second-order effects. The results were complemented with 95% confidence intervals, allowing for a more rigorous interpretation of the influence each factor has on the variability of the obtained results.

Table 5 presents the  $S_1$ , ST and half-width (HW) of the 95% confidence interval. MW emerges as the most relevant parameter in both scenarios, with  $S_1$  values of 0.643 in dilution and 0.533 in overbreak. This suggests that adjusting the MW can significantly reduce mineral losses and unwanted excavations. On the other hand, both PF and EQ exert a notable influence, especially in overbreak situations, with  $S_1$  values of 0.129 and 0.102, respectively. This highlights that an

inadequate configuration of explosive energy significantly contributes to overbreak. Therefore, optimizing fragmentation through a proper selection of specific energy can positively impact stope stability. Regarding RMR, its impact is more significant in dilution ( $S_1=0.115$ ) than in overbreak ( $S_1=0.057$ ), reinforcing the idea that weaker rocks experience higher mineral losses. The discrepancy between  $S_1$  and ST in variables such as PF, EQ, and AL suggests that their effect is not only direct but also depends on interactions with other factors. Hence, controlling these parameters must be approached holistically rather than in isolation. Finally, the minimal impact of CL and CF indicates that managing explosive charge design should prioritize the total amount of explosive and its distribution, rather than focusing solely on CL.

The  $S_2$  interaction analysis between key variables is presented in Table 6. The results indicate that RMR has a significant influence on the MW, as the coefficients remain stable at 0.054 for overbreak and 0.052 for dilution, suggesting that lower rock quality amplifies this effect. The powder factor (PF) also affects MW, with coefficients of 0.043 and 0.033, indicating that excessive explosive energy results in unnecessary fracturing, while an insufficient charge compromises rock fragmentation. The EQ must also be adjusted, as coefficients of 0.051 for overbreak and 0.055 for dilution indicate that an overload leads to excessive excavation, whereas an insufficient charge leaves unblasted rock. Finally, the EQ-RMR interaction confirms that rock mass strength determines the required explosive charge, as harder rocks demand more energy, whereas in weaker rock masses, excessive explosives increase dilution and overbreak.

**Table 5.**  $S_1$  and ST values

Target	Variable	$S_1$	$S_1$ -HW	ST	ST-HW
Dilution	MW	0.643	0.028	0.681	0.019
	RMR	0.115	0.014	0.152	0.006
	PF	0.049	0.012	0.092	0.004
	AL	0.045	0.011	0.078	0.004
	EQ	0.040	0.012	0.080	0.003
	RQD	0.033	0.011	0.069	0.003
	CF	0.026	0.010	0.065	0.003
	CL	0.008	0.009	0.050	0.002
Overbreak	MW	0.533	0.028	0.573	0.022
	PF	0.129	0.015	0.177	0.007
	EQ	0.102	0.013	0.147	0.007
	RMR	0.057	0.013	0.099	0.004
	AL	0.053	0.011	0.087	0.004
	CL	0.043	0.012	0.090	0.004
	RQD	0.039	0.010	0.077	0.003
	CF	0.010	0.008	0.052	0.003

**Table 6.**  $S_2$  values

Variables	$S_2$ -Overbreak	$S_2$ -Dilution
MW-RMR	0.054	0.052
MW-PF	0.043	0.033
MW-EQ	0.051	0.055
EQ-RMR	0.042	0.057

Therefore, to reduce dilution and overbreak, optimization should focus on ensuring that the MW is adjusted to optimal operational limits, guaranteeing both stability and safety. The controllability of explosive energy must also be considered, as adjusting the powder factor and the amount of explosive should help reduce unwanted excavation and operational risks.

Proper characterization of the rock mass before designing drilling and blasting is essential to prevent structural failures and improve efficiency. Finally, optimizing PF, EQ, and considering RMR will lead to a more precise, efficient, and safer excavation.

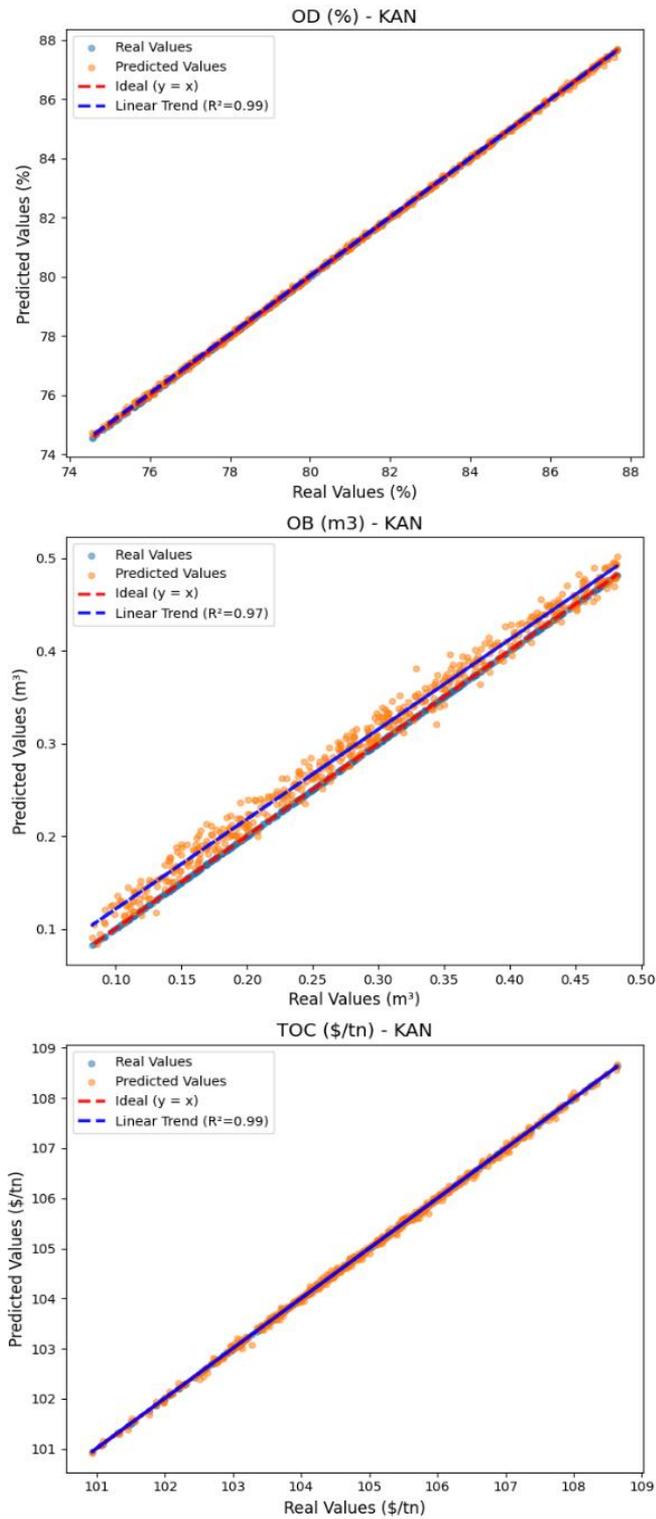


Figure 9. Prediction of OD (%)-OB(m<sup>3</sup>) – TOC (\$/ton)

### 3.4 Application in an underground mine

The KAN model was applied to an underground mine located in the province of Pataz, northern Peru. The mining operation shares similar characteristics with the variables used

to train the model, ensuring its applicability. The dataset used for testing consisted of 246 records.

Figure 9 indicates that the model demonstrates strong predictive potential for dilution, with a coefficient of determination of R<sup>2</sup>=0.99. Similarly, its prediction accuracy for overbreak is also high, with R<sup>2</sup>=0.97. Additionally, the model exhibits good predictive performance in estimating operational costs, reinforcing its effectiveness in optimizing the mining process.

Once its validity was verified, its effectiveness was compared with traditional methods (TM) used in this mine to estimate dilution and overbreak. In this mine, the operational dilution estimation method considers the impact on the footwall and hanging wall due to overbreak caused by explosive detonation, taking into account the mining section, blast effects, geomechanical conditions, and the cleaning of extracted ore. This combination of factors determines the designed operational dilution percentage, which is calculated using the formulas established for this case by Pakalnis. To improve overbreak estimation, numerical models such as the Finite Element Method (FEM) are used.

A total of 30 blasts were evaluated, comparing the performance of the KAN model against traditional methods. Figure 10 shows that KAN achieved a 97% accuracy in predicting dilution, significantly outperforming the traditional method, which obtained 77%.

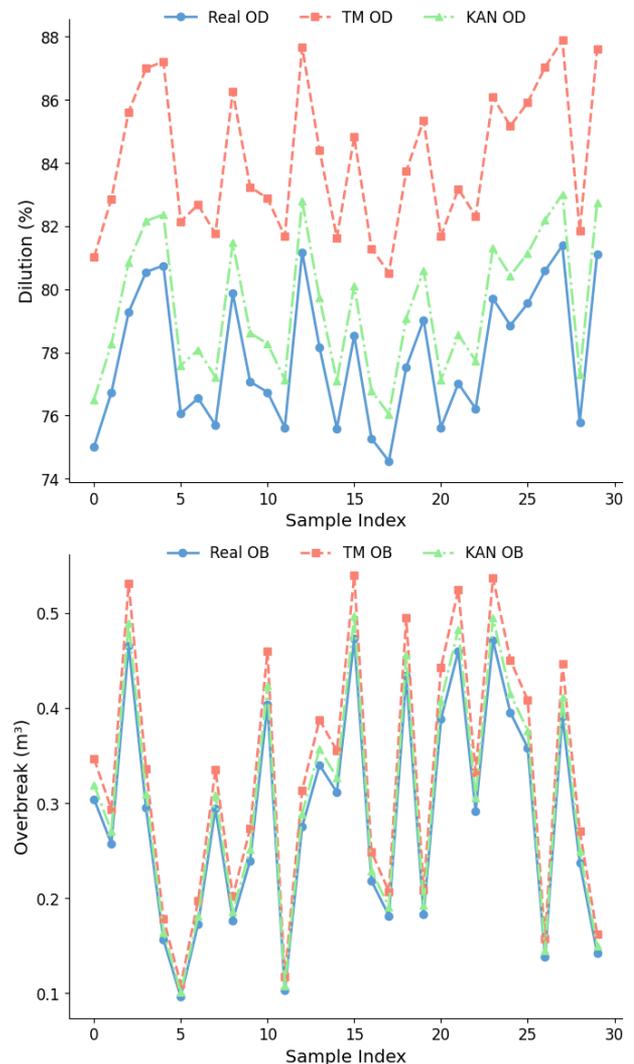


Figure 10. Comparison real vs. TM vs. KAN

Regarding overbreak, the KAN model reached an accuracy of 94%, whereas the FEM achieved 86%, demonstrating the superiority of KAN in estimating these key parameters for optimizing mine design.

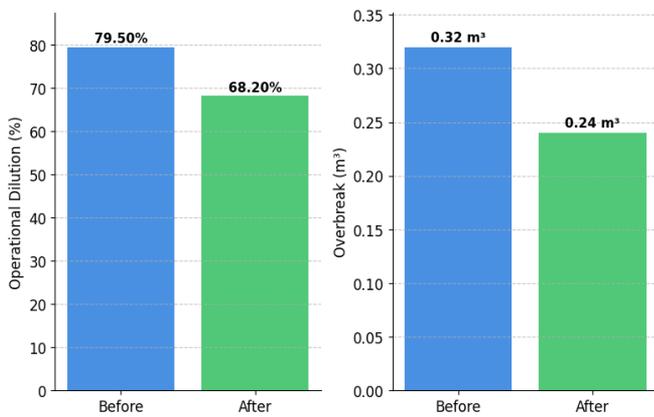
### 3.5 Parameter optimization with MOPSO

MOPSO was used exclusively to determine the optimal controllable parameters, such as MW (m), CL (m), AL (m), and EQ (kg). This approach focuses on optimization on directly adjustable variables to minimize dilution, overbreak, and costs.

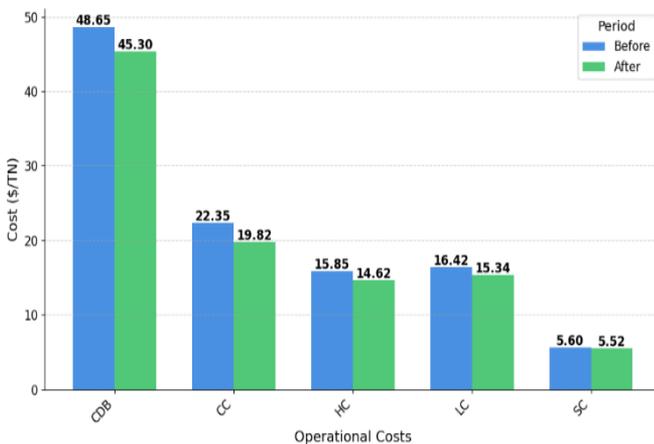
Table 7 presents the optimal values identified by MOPSO. The MW shows a difference of 3.38%, which remains within an acceptable margin, ensuring consistency in mining practices. The CL deviates by 3.94%, indicating potential opportunities to optimize material loading precision. Similarly, the AL exhibits a variation of 2.76%, which, while within operational limits, could be further refined to enhance efficiency. Finally, the quantity of explosive used exceeds the optimal value by 5.07%.

**Table 7.** Main optimized parameters

Parameter	Optimal Value	Real Value	Difference (%)
MW (m)	2.30	2.38	3.38%
CL (m)	1.27	1.32	3.94%
AL (m)	1.45	1.49	2.76%
EQ (kg)	14.60	15.34	5.07%



**Figure 11.** OD (%) - OB(m³) before vs. after



**Figure 12.** Operational costs before vs. after

Although these variables allow for a high degree of control, their adjustment is constrained by technical limitations, operating costs, safety regulations, and local geomechanically conditions. In this case, the RMR averages 65 (IIB), classifying the rock as good quality.

The optimal values recommended for use in drilling and blasting design were analyzed according to the geomechanically factors of the operation, achieving changes that improve rock mass stability by reducing overbreak and minimizing dilution. These adjustments can lead to increased operational efficiency, better material fragmentation, and optimized explosive consumption, thereby ensuring the safety and profitability of the mining process.

Figure 11 presents a comparison between the average values from the database and the improvements achieved by the model after applying the new parameters in the simulation. The initial average dilution was 79.5%, whereas the model successfully reduced it to 68.2%, decreasing by 11.3%. This represents a significant improvement in the accuracy of dilution control. Similarly, for overbreak, which initially averaged 0.32 m³, the model optimized it to 0.24 m³, decreasing by 25%, indicating a notable reduction in unwanted material volumes.

Figure 12 shows that for drilling and blasting, the cost decreased from 48.65 to 45.30 US\$/Tn, representing a 6.89% reduction. In clean-up, the cost dropped from 22.35 to 19.82 US\$/Tn, equating to an 11.33% saving. Haulage costs declined from 15.85 US\$/Tn to 14.62 US\$/Tn, showing a 7.77% improvement, while loading costs decreased from 16.42 to 15.34 US\$/Tn, resulting in a 6.58% saving. Finally, the support cost was reduced from 5.60 to 5.52 US\$/Tn, achieving a 1.43% reduction.

Overall, the total operating cost decreased from 108.87 to 100.60 US\$/Tn, representing a global optimization of 7.60%. These improvements support the synergy between KAN and MOPSO, highlighting how their combination enhances modelling capabilities by providing more accurate and applicable solutions in the mining sector. This approach establishes itself as a solid and efficient framework for addressing optimization in complex underground mining scenarios.

Chimunhu et al. [15] utilized geological, geotechnical, and design data accessible early in the mine planning phase to forecast dilution, attaining a remarkable accuracy of 93% with the PCA-CART hybrid model. Their investigation also observed a minor underestimation of dilution by 13% relative to the stopes in the initial design. In a similar vein, Jang et al. [17] applied a neuro-fuzzy system to predict unplanned dilution and ore loss, achieving an R-value of 0.72. Their model also offered practical recommendations to reduce dilution and ore loss, emphasizing the importance of factors like powder factor and ground support, highlights the critical role of selecting the right variables.

This study included significant geometric aspects, such as MW, and geomechanically characteristics like RMR, in addition to drilling and blasting variables, to guarantee precise forecasts of dilution and overbreak. Hong et al. [22] determined that tunnel diameter was the predominant factor in calculating overbreak, accounting for 90.5%, with RMR contributing 1.80%. Their BO-XGBoost model attained a robust R² score of 0.94. Jang and Topal [18] utilized RMR and overbreak as input and output parameters, respectively, attaining a remarkable R² of 0.95 with an optimized ANN. Mottahedi et al. [24] conducted a sensitivity analysis utilizing

the cosine amplitude method (CAM) and found that the S/B and S/D parameters exerted the most significant influence on their model, whereas RMR demonstrated an effectiveness of 0.79 in the ANFIS-PSO model, achieving an  $R^2$  of 0.96. Similarly, this study showed that MW is the most influential variable responsible for overbreak and dilution. It contributes 64.3% to dilution and 53.3% to overbreak, also highlighting the importance of RMR and RQD in the prediction of these targets, achieving an  $R^2$  greater than 0.95.

Guo et al. [27] integrated the MOPSO algorithm into their model, thereby identifying the optimal set of blasting parameters. In practical operational environments, these parameters demonstrated both efficacy and reliability, resulting in a 3% decrease in total costs across the mining process, encompassing drilling, blasting, loading, hauling, and crushing. This research demonstrated that MOPSO facilitated the KAN model in precisely determining the parameters to prioritize for minimizing dilution and overbreak, achieving a 7.60% reduction in total operating expenses.

## 4. CONCLUSIONS

### 4.1 Key findings, importance, and relevance

In this research, dilution, overbreak, and costs were successfully predicted using KAN networks in underground mining.

The KAN model was implemented, and its effectiveness was assessed in comparison to the MLP, RF, and XGBoost models. The results demonstrated that KAN outperformed the other models in forecasting dilution, with metrics of MSE=0.001, RMSE=0.032, MAE=0.055, and  $R^2=0.99$ , indicating an exceptional match. For overbreak, KAN achieved MSE=0.00, RMSE=0.00, MAE=0.008, and  $R^2=0.99$ . For costs, KAN exhibited MSE=0.001, RMSE=0.032, MAE=0.027, and  $R^2=0.99$ , signifying exceptional performance relative to other evaluated models.

Sobol sensitivity analysis on KAN model showed that MW is the most influential variable responsible for overbreak and dilution. It contributes 64.3% to dilution and 53.3% to overbreak, thus confirming its predominant influence in these findings. Other moderate contributory factors of overbreak are EQ and PF, which are EQ and powder factor, respectively, contributing about 10.2% and 12.9%.

The MOPSO algorithm was used to optimize the drilling and blasting parameters based on design specifications. The model indicated the optimal parameters: an MW of 2.30 m, an AL of 1.27 m, a CL of 1.45 m, and 14.60 kg of explosives. When analyzed and implemented, these parameters resulted in an 11.3% reduction in operational dilution, a 25% decrease in overbreak, and a 7.60% reduction in total operating costs.

These findings represent a crucial advancement in predictive modeling for underground mining, providing practical solutions to genuinely enhance operational efficiency while ensuring operational safety and profitability in the mining process.

### 4.2 Limitations of the study and recommendations for future work

The KAN has proven to be highly robust in predicting dilution, overbreak, and operational costs. However, its effectiveness relies on the careful selection and use of spline

functions, which are crucial for accurately modeling the nonlinear relationships between input and output variables. These findings strongly depend on the quality and representativeness of the dataset, which may limit their applicability to other mining contexts. This issue becomes particularly critical due to the various parameters that influence different mining environments. Therefore, there is a need to expand the analysis and modify the model to adapt to more varied and complex conditions.

Future research should focus on automating and refining the selection and adaptation of basic functions through approaches such as evolutionary algorithms or Bayesian optimization. These enhancements would improve the KAN model's ability to learn nonlinear patterns while mitigating overfitting. Additionally, conducting a comprehensive analysis of the variables and parameters included in the model is recommended. Ultimately, expanding the dataset and validating the model across various mining scenarios with diverse attributes would enhance its generalization and ensure broader applicability.

## REFERENCES

- [1] Al-Albakri, A., Hefni, M. (2021). A review of some nonexplosive alternative methods to conventional rock blasting. *Open Geosciences*, 13(1): 431-442. <https://doi.org/10.1515/geo-2020-0245>
- [2] Yilmaz, O. (2023). Drilling and blasting designs for parallel hole cut and V-cut method in excavation of underground coal mine galleries. *Scientific Reports*, 13(1): 2449. <https://doi.org/10.1038/s41598-023-29803-6>
- [3] Conde, F., Sanoh, O. (2022). Analysis and optimization of blasting practices at the Sangaredi mine. *Journal of Geoscience and Environment Protection*, 10(9): 149-169. <https://doi.org/10.4236/gep.2022.109010>
- [4] Verma, H.K., Samadhiya, N.K., Singh, M., Goel, R.K., Singh, P.K. (2018). Blast induced rock mass damage around tunnels. *Tunnelling and Underground Space Technology*, 71: 149-158. <https://doi.org/10.1016/j.tust.2017.08.019>
- [5] Cordova, D.P., Zingano, A.C., Gonçalves, Í.G. (2022). Unplanned dilution back analysis in an underground mine using numerical models. *REM - International Engineering Journal*, 75(4): 379-388. <https://doi.org/10.1590/0370-44672021750093>
- [6] Mussin, A., Kydrashov, A., Asanova, Z., Abdrakhman, Y., Ivdilinova, D. (2024). Ore dilution control when mining low-thickness ore bodies using a system of sublevel drifts. *Mining of Mineral Deposits*, 18(2): 18-27. <https://doi.org/10.33271/mining18.02.018>
- [7] Câmara, T.R., Leal, R.S., de Lemos Peroni, R., Capponi, L.N. (2019). Controlling operational dilution in open-pit mining. *Mining Technology*, 128(1): 1-8. <https://doi.org/10.1080/25726668.2018.1470275>
- [8] Azaryan, A.A., Batareyev, O.S., Karamanits, F.I., Kolosov, V.O., Morkun, V.S. (2024). Ways to reduce ore losses and dilution in iron ore underground mining in Kryvbass. *Science and Innovation*, 14(4): 17-24. <https://doi.org/10.15407/scine14.03.017>
- [9] Montalvo, M., Ordóñez, C., Reyes, E. (2024). Minimization of Ore dilution in the "Pique" mine through pre-splitting blasting technique. *ESPOCH*

- Congresses: The Ecuadorian Journal of STEAM, 3(2): 244-261. <https://doi.org/10.18502/epoch.v4i1.15828>
- [10] Rodovalho, E., El Hajj, T.M., Teodoro, G.C., de Tomi, G., Tenório, J.A.S. (2024). Reducing mining tailings and operational dilution: A new application of the room-and-pillar mining method. *Journal of the Southern African Institute of Mining and Metallurgy*, 124(8): 483-489. <https://doi.org/10.17159/2411-9717/1696/2024>
- [11] Patson, V., Maxmudov, D., Nurboboiev, Y., Mahmadioliev, A., Pardaev, F. (2021). Impact of ore dilution on the technical and economic performance of a mining enterprise. *Technical Science and Innovation*, 1: 149-155.
- [12] Chimunhu, P., Topal, E., Asad, M., Faradonbeh, R., Ajak, A. (2024). The future of underground mine planning in the era of machine learning: Opportunities for engineering robustness and flexibility. *Mining Technology*, 133(4): 331-347. <https://doi.org/10.1177/25726668241281875>
- [13] Chimunhu, P., Topal, E., Ajak, A., Asad, W. (2022). A review of machine learning applications for underground mine planning and scheduling. *Resources Policy*, 77: 102693. <https://doi.org/10.1016/j.resourpol.2022.102693>
- [14] Jorquera, M., Korzeniowski, W., Skrzypkowski, K. (2023). Predicción de la dilución en el método de explotación por sublevel stoping mediante algoritmos de aprendizaje automático. *IOP Conference Series: Earth and Environmental Science*, 1189(1): 012008. <https://doi.org/10.1088/1755-1315/1189/1/012008>
- [15] Chimunhu, P., Faradonbeh, R., Topal, E., Mohamed Waqar, A., Ajak Duany, A. (2024). Development of novel hybrid intelligent predictive models for dilution prediction in underground sub-level mining. *Mining, Metallurgy & Exploration*, 41(4): 2079-2098. <https://doi.org/10.1007/s42461-024-01029-8>
- [16] Zhao, X., Niu, J. (2020). Method of predicting ore dilution based on a neural network and its application. *Sustainability*, 12(4): 1550. <https://doi.org/10.3390/su12041550>
- [17] Jang, H., Topal, E., Kawamura, Y. (2015). Decision support system of unplanned dilution and ore-loss in underground stoping operations using a neuro-fuzzy system. *Applied Soft Computing*, 32: 1-12. <https://doi.org/10.1016/j.asoc.2015.03.043>
- [18] Jang, H., Topal, E. (2013). Optimizing overbreak prediction based on geological parameters comparing multiple regression analysis and artificial neural network. *Tunnelling and Underground Space Technology*, 38: 161-169. <https://doi.org/10.1016/j.tust.2013.06.003>
- [19] Koopialipoor, M., Jahed Armaghani, D., Haghghi, M., Noroozi Ghaleini, E. (2019). A neuro-genetic predictive model to approximate overbreak induced by drilling and blasting operation in tunnels. *Bulletin of Engineering Geology and the Environment*, 78: 981-990. <https://doi.org/10.1007/s10064-017-1116-2>
- [20] Koopialipoor, M., Noroozi Ghaleini, E., Haghghi, M., Kanagarajan, S., Maarefvand, P., Tonnizam Mohamad, E. (2019). Overbreak prediction and optimization in tunnel using neural network and bee colony techniques. *Engineering with Computers*, 35: 1191-1202. <https://doi.org/10.1007/s00366-018-0658-7>
- [21] Koopialipoor, M., Noroozi Ghaleini, E., Tootoonchi, H., Jahed Armaghani, D., Haghghi, M., Hedayat, A. (2019). Developing a new intelligent technique to predict overbreak in tunnels using an artificial bee colony-based ANN. *Environmental Earth Sciences*, 78: 165. <https://doi.org/10.1007/s12665-019-8163-x>
- [22] Hong, Z., Tao, M., Liu, L., Zhao, M., Wu, C. (2023). An intelligent approach for predicting overbreak in underground blasting operation based on an optimized XGBoost model. *Engineering Applications of Artificial Intelligence*, 126: 107097. <https://doi.org/10.1016/j.engappai.2023.107097>
- [23] Mottahedi, A., Sereshki, F., Ataei, M. (2018). Development of overbreak prediction models in drill and blast tunneling using soft computing methods. *Engineering with Computers*, 34: 45-58. <https://doi.org/10.1007/s00366-017-0520-3>
- [24] Mottahedi, A., Sereshki, F., Ataei, M. (2018). Overbreak prediction in underground excavations using hybrid ANFIS-PSO model. *Tunnelling and Underground Space Technology*, 80: 1-9. <https://doi.org/10.1016/j.tust.2018.05.023>
- [25] Bakhtavar, E., Abdollahisharif, J., Ahmadi, M. (2016). Reduction of the undesirable bench-blasting consequences with emphasis on ground vibration using a developed multi-objective stochastic programming. *International Journal of Mining, Reclamation and Environment*, 31(5): 333-345. <https://doi.org/10.1080/17480930.2016.1158964>
- [26] Guo, J., Zhao, P., Li, P. (2023). Prediction and optimization of blasting-induced ground vibration in open-pit mines using intelligent algorithms. *Applied Sciences*, 13(12): 7166. <https://doi.org/10.3390/app13127166>
- [27] Guo, J., Zhao, Z., Zhao, P., Chen, J. (2024). Prediction and optimization of open-pit mine blasting based on intelligent algorithms. *Applied Sciences*, 14(13): 5609. <https://doi.org/10.3390/app14135609>
- [28] Sharma, V. (2022). A study on data scaling methods for machine learning. *International Journal for Global Academic & Scientific Research*, 1(1): 31-42. <https://doi.org/10.55938/ijgasr.v1i1.4>
- [29] Gorriz, J.M., Segovia, F., Ramirez, J., Ortiz, A., Suckling, J. (2024). Is K-fold cross validation the best model selection method for Machine Learning? *arXiv preprint arXiv:2401.16407*. <https://doi.org/10.48550/arXiv.2401.16407>
- [30] Refaeilzadeh, P., Tang, L., Liu, H. (2009). Cross-validation. In *Encyclopedia of Database Systems*, pp. 532-538. [https://doi.org/10.1007/978-0-387-39940-9\\_565](https://doi.org/10.1007/978-0-387-39940-9_565)
- [31] Liu, Z., Wang, Y., Vaidya, S., Ruehle, F., et al. (2024). Kan: Kolmogorov-Arnold networks. *arXiv preprint arXiv:2404.19756*. <https://doi.org/10.48550/arXiv.2404.19756>
- [32] Khlef, A. (2024). A comprehensive review of applications of spline function. *International Academic Journal of Science and Engineering*, 11(1): 322-331. <https://doi.org/10.9756/IAJSE/V11I1/IAJSE1138>
- [33] Jamali, A., Swalpa Kumar, R., Danfeng, H., Bing, L., Pedram, G. (2024). How to learn more? Exploring Kolmogorov-Arnold Networks for hyperspectral image classification. *Remote Sensing*, 16(21): 4015. <https://doi.org/10.3390/rs16214015>
- [34] Xu, K., Chen, L., Wang, S. (2024). Kolmogorov-Arnold

- networks for time series: Bridging predictive power and interpretability. arXiv preprint arXiv:2406.02496. <https://doi.org/10.48550/arXiv.2406.02496>
- [35] Bochra, K. (2024). Kolmogorov-Arnold networks: Key developments and uses. *Qeios*. <https://doi.org/10.32388/7NNCAA>
- [36] Ibrahim, A.D.M., Shang, Z., Hong, J.E. (2024). How resilient are Kolmogorov–Arnold networks in classification tasks? A robustness investigation. *Applied Sciences*, 14(22): 10173. <https://doi.org/10.3390/app142210173>
- [37] Ta, H.T. (2024). BSRBF-KAN: A combination of B-splines and radial basis functions in Kolmogorov-Arnold networks. arXiv preprint arXiv:2406.11173. <https://doi.org/10.48550/arXiv.2406.11173>
- [38] Sasse, A.M., de Farias, C.M. (2024). Evaluating federated Kolmogorov-Arnold networks on non-IID data. arXiv preprint arXiv:2410.08961. <https://doi.org/10.48550/arXiv.2410.08961>
- [39] Shukla, K., Toscano, J.D., Wang, Z., Zou, Z., Karniadakis, G.E. (2024). A comprehensive and fair comparison between MLP and KAN representations for differential equations and operator networks. *Computer Methods in Applied Mechanics and Engineering*, 431: 117290. <https://doi.org/10.1016/j.cma.2024.117290>
- [40] Saillot, M., Dominique, M., Ahmed, Z. (2024). B-spline curve approximation with transformer neural networks. *Mathematics and Computers in Simulation*, 223: 275-287. <https://doi.org/10.1016/j.matcom.2024.04.010>
- [41] Kotsiopoulos, T., Sarigiannidis, P., Ioannidis, D., Tzovaras, D. (2021). Machine learning and deep learning in smart manufacturing: The smart grid paradigm. *Computer Science Review*, 40: 100341. <https://doi.org/10.1016/j.cosrev.2020.100341>
- [42] Cabanilla, K., Mohammad, R., Lope, J. (2024). Neural networks with ReLU powers need less depth. *Neural Networks*, 172: 106073. <https://doi.org/10.1016/j.neunet.2023.12.027>
- [43] Kingma, D.P., Ba, J. (2014). Adam: A method for stochastic optimization. arXiv preprint arXiv:1412.6980. <https://doi.org/10.48550/arXiv.1412.6980>
- [44] Salman, H., Kalakech, A., Steiti, A. (2024). Random forest algorithm overview. *Babylonian Journal of Machine Learning*, 2024: 69-79. <https://doi.org/10.58496/BJML/2024/007>
- [45] Schauburger, G., Klug, S.J., Berger, M. (2024). Random forests for the analysis of matched case–control studies. *BMC Bioinformatics*, 25(1): 253. <https://doi.org/10.1186/s12859-024-05877-5>
- [46] Belyadi, H., Haghghat, A. (2021). Chapter 5 - Supervised learning. In *Machine Learning Guide for Oil and Gas Using Python*, pp. 169-295. <https://doi.org/10.1016/B978-0-12-821929-4.00004-4>
- [47] Chen, T., Guestrin, C. (2016). XGBoost: A scalable tree boosting system. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, San Francisco, USA, pp. 785-794. <https://doi.org/10.1145/2939672.2939785>
- [48] Karunasingha, D. (2022). Root mean square error or mean absolute error? Use their ratio as well. *Information Sciences*, 585: 609-629. <https://doi.org/10.1016/j.ins.2021.11.036>
- [49] Gao, J. (2024). R-Squared ( $R^2$ ) - How much variation is explained? *Research Methods in Medicine & Health Sciences*, 5(4): 104-109. <https://doi.org/10.1177/26320843231186398>