Rock Mechanics and Rock Engineering (2025) 58:679–695
https://doi.org/10.1007/s00603-024-04166-0

ORIGINAL PAPER

Adaptive Weighted Multi-kernel Learning for Blast-Induced Flyrock Distance Prediction

Ruixuan Zhang[1] · Yuefeng Li[2] · Yilin Gui[1,3,4] · Danial Jahed Armaghani[5,6] · Mojtaba Yari[7]

Received: 7 August 2023 / Accepted: 5 September 2024 / Published online: 25 September 2024

Abstract

In the field of civil and mining engineering, blasting operations are widely and frequently used for rock excavation, However, some undesirable environmental problems induced by blasting operations cannot be ignored. Blast-induced flyrock is one important issue induced by blasting operation, which needs to be well predicted to identify the blasting zone's safety zone. This study introduces an adaptive weighted multi-kernel learning model (AW-MKL) to provide an accurate prediction of blast-induced flyrock distance in Sungun Copper Mine site. The proposed model uses a combination of multi-kernel learning (MKL) approach and adaptive weighting strategy based on weighted Euclidean distance and modified local outlier factor (MLOF) to maximally improve the predictive ability of kernel ridge regression (KRR). To demonstrate the superiority of the proposed approach, six machine learning models were developed as comparisons, i.e., KRR, RF, GBDT, SVM, M5 Tree, MARS and AdaBoost. The outcomes of the proposed method achieved the highest accuracy in testing phase, with RMSE of 2.05, MAE of 0.98 and VAF of 99.92, which confirmed the strong predictive capability of the proposed AW-MKL in predicting blast-induced flyrock distance.

Highlights

- An adaptive weighted multi-kernel learning model (AW-MKL) to predict blast induced flyrock.
- Improve the performance of kernel ridge regression (KRR) using the combination of multi-kernel learning (MKL) and adaptive weighting.
- Use MKL to reduce the effort devoted in determine optimal kernel.
- Measure the correlation between model input and training samples using weighted Euclidean distance.
- Propose a modified local outlier factor (MLOF) to evaluate the uncertainty of training samples.
- The proposed adaptive weighting strategy assign weights to training samples according to the weighted Euclidean distance and MLOF.

Keywords Rock blasting · Flyrock · Multiple kernel learning · Modified local outlier factor · Adaptive weighting

✉ Yilin Gui
yilin.gui@qut.edu.au

[1] School of Civil and Environmental Engineering, Queensland University of Technology, Gardens Point, Brisbane, QLD 4000, Australia

[2] School of Computer Science, Queensland University of Technology, Gardens Point, Brisbane, QLD 4000, Australia

[3] Centre for Materials Science, Queensland University of Technology, Gardens Point, Brisbane, QLD 4000, Australia

[4] Centre for Sustainable Engineered Construction Materials, Queensland University of Technology, Gardens Point, Brisbane, QLD 4000, Australia

[5] Centre of Tropical Geoengineering (GEOTROPIK), Institute of Smart Infrastructure and Innovative Engineering (ISIIC), Faculty of Civil Engineering, Universiti Teknologi Malaysia, 81310 Johor Bahru, Malaysia

[6] Department of Urban Planning, Engineering Networks and Systems, Institute of Architecture and Construction, South Ural State University, 76 Lenin Prospect, Chelyabinsk, Russia 454080

[7] Department of Mining Engineering, Faculty of Engineering, Malayer University, Malayer 65719-95863, Iran

1 Introduction

Blasting operations are the most conventional method to produce desired rock fragmentation in civil and mining construction works (Bhandari 1997). However, only 20–30% of the explosive energy generated by explosives can be used for rockmass breakage, and the remaining part is wasted to surrounding environment, leading to undesirable consequences like ground vibration, airblast and flyrock (Yari et al. 2017; Harandizadeh and Armaghani 2021; Murlidhar et al. 2021; Monjezi et al. 2010, 2011). In open pit mines, flyrock is considered as the main hazard of blasting operations, because it may cause fatal or non-fatal events (Bajpayee et al. 1999). The blast-induced flyrock phenomenon is defined as the broken rock particles boosted far from the free face after blasting operations (Yari et al. 2014). Factors affecting flyrock can be concluded into two categories, controllable factors and uncontrollable factors, and both of them should be considered to design a reliable blasting operation (Yari et al. 2013; Jahed Armaghani et al. 2016). For controllable factors, flyrock is mainly affected by drilling, blasthole diameter, specific charge, stemming, delay time and burden (Rezaei et al. 2011). Inappropriate geologic features and loosing rock at the top of free face are the main uncontrollable sources that may lead to flyrock (Raina et al. 2004).

To predict flyrock distance induced by blasting operations, traditional studies used empirical equations to estimate the relationship between flyrock distance and its influencing factors. One famous empirical equation was proposed by Lundborg et al. (1975), using the hole diameter as variable for flyrock distance estimation. The relationship between stemming to burden ratio and flyrock distance was explored in the study (Olofsson 1990). An empirical equation considering the effect of both uncontrollable parameters (rock quality designation) and controllable parameter (stemming to burden ratio) was developed by Trivedi et al. (2014) to predict flyrock distance. However, the accuracy of empirical equations is limited due to several limitations like the diversity of influencing factors and the inherent uncertainty of blasting operations (Hustrulid 1999; Monjezi et al. 2007).

Considering the limitations of empirical equations, the machine learning based predictive approach was considered by many researchers for flyrock distance prediction due to its strong non-linear approximation ability. As a traditional and widely used machine learning method, artificial neural network (ANN) models were developed by many studies to predict blast-induced flyrock distance (Trivedi et al. 2014, 2015; Ghasemi et al. 2014; Mohamad et al. 2013; Monjezi et al. 2013; Saghatforoush et al. 2016; Guo et al. 2021). In the studies (Jahed Armaghani et al. 2016; Rezaei et al. 2011), flyrock distance predictive model were designed using fuzzy inference system (FIS) and its variant, adaptive neuro fuzzy inference system (ANFIS) respectively. Tree based models were also implemented for flyrock distance estimation. The basic classification and regression tree (CART) model was trained by Hasanipanah et al. (2017) using 65 sets of blasting data collected from Ulu Tiram quarry, Malaysia. The ensemble tree models, like random forest (RF) and extreme gradient boosting (XGBoost) were developed by Han et al. (2020) and Yari et al. (2023), showing that they are capable for flyrock distance estimation. Other than these models, support vector regression (SVR) were also available for prediction of blast-induced flyrock distance (Amini et al. 2012; Rad et al. 2018). In these studies, the developed machine learning models have shown their superiority over empirical equations by achieving higher accuracies. However, their approximation abilities are still limited because they suffer from various limitations. For example, the ANN has the problem of local minima, and the performance of SVR is highly depend on the hyperparameters. Therefore, hybrid models were developed by many studies to address this issue where meta-heuristic algorithms were used to further improve the performance of machine learning models through an iterative optimization process. Marto et al. (2014), Monjezi et al. (2012) and Hasanipanah et al. (2022) proved the local minima problem of ANN can be solved using imperialist competitive algorithm (ICA), genetic algorithm (GA) and adaptive dynamical harmony search algorithm. The performances of ICA, GA and particle swarm optimization (PSO) in optimizing ANN were compared in the study (Koopialipoor et al. 2019). Nguyen et al. (2021) used whale optimization algorithm (WOA) to tune the hyperparameters of SVR under different kernel functions. Although hybrid models are able to achieve better predictive performances, its drawback is obvious, that is time-consuming. This is because the meta-heuristic algorithm optimizes the model through an iterative process where the model always needs to be trained thousands of times to achieve its optimum performance.

In this study, an attempt had been made to develop an adaptive-weighted multi-kernel learning method (AW-MKL). The predictive power of KRR has already been examined in the field of predicting blast-induced flyrock distance by Jamei et al. (2021). Hence, this study attempts to further improve the performance of KRR by introducing the adaptive weighting strategy and multiple kernel learning (MKL) method. The AW-MKL is proposed considering two aspects. Initially, considering the size of dataset, reweighting of training samples according to the variation of input can be applicable to improve the model performance since the computational cost is not huge. Moreover, the development of KRR requires careful selection on kernel function and its parameters to achieve the optimal performance. Most of the existing studies used meta-heuristic algorithms for hyperparameters tuning (Ding et al. 2024; Chen et al. 2022), which is time consuming (Zhang et al.

2022). Hence, the multi-kernel learning method is applicable for reducing the effort devoted in tuning the optimal kernel by developing a composite kernel function. Considering these two issues, the combination of reweighting strategy and MKL method is applied to improve the predictive power of KRR method. An existing study has proved the applicability of reweighted KRR model (Zhang et al. 2023), and the proposed AW-MKL extended it into multiple kernel version. Beyond that, the proposed AW-MKL reallocates the model attention considering pairwise similarity as well as the uncertainty degree of training sample. For this issue, a weighted Euclidean distance is used to measure the similarity between samples. Using the similarity based weighting strategy, those training samples with high similarity can contribute more to the model output. Furthermore, considering the uncertainties involved in blasting operations (Ghasemi et al. 2012; Armaghani et al. 2016), some datapoints might violate the general relationship between flyrock distance and its influencing factors and are therefore hard to be fitted. These datapoints might generate negative impact on the overall fitting performances of the model. Hence, a modified local outlier factor (MLOF) is proposed to address this issue, and used with the aforementioned weighted Euclidean distance to weaken the impact of those datapoints with higher uncertainties. The combination of weighted Euclidean distance and MLOF based sample similarity weighting strategy is called adaptive weighting in this study, which is applied to the MKL model to further improve the predictive capability and reliability of MKL model. To examine the reliability of the proposed method, seven other machine learning models were implemented as comparison, i.e., KRR, SVM, RF, GBDT, MARS, M5 Tree and AdaBoost.

This paper is organized as following. Section 2 describes the theoretical concepts of the proposed AW-MKL. In Sect. 3, the details of the dataset used and indices used for model evaluation are introduced. Then, the developments of seven predictive models are also presented in this section. Their obtained results and a corresponding discussion are reported in the following Sect. 4. In the last section, conclusions are drawn including the main results, contributions and limitations of the proposed model.

2 Background

2.1 Weighted Multi-kernel Learning

2.1.1 Weighted Kernel Ridge Regression

Kernel ridge regression (KRR) was proposed in 1998 (Saunders et al. 1998). It is the nonlinear version of ridge regression using kernel theory. In this study, we perform multi-input and single-output regression case and the details of KRR are therefore presented as following. Denote the training samples as $D = \{\boldsymbol{x}_i, y_i\}_{i=1}^{n}$, where $\boldsymbol{x}_i \in \mathbb{R}^M$ is the M-dimensional model input and y_i represents the desired output. In linear regression case, the output is approximated by $y = w \cdot x$, and the objective function of ridge regression can be expressed as:

$$\arg\min : \frac{C}{2}\|\boldsymbol{Y} - \boldsymbol{X}\boldsymbol{w}\|^2 + \frac{1}{2}\|\boldsymbol{w}\|^2 \tag{1}$$

where $\boldsymbol{X} = \begin{bmatrix} \boldsymbol{x}_1 \\ \vdots \\ \boldsymbol{x}_n \end{bmatrix}$, $\boldsymbol{w} = \begin{bmatrix} w_1 & \cdots & w_n \end{bmatrix}^{\mathrm{T}}$ and $\boldsymbol{Y} = \begin{bmatrix} y_1 & \cdots & y_n \end{bmatrix}^{\mathrm{T}}$.

The right-hand term in the above equation is a regularization term, which is used to minimize the structural risk. C is a positive constant controlling the degree to which the model tends to minimize the training error or the structural risk.

To perform nonlinear regression, the training samples are mapped through a nonlinear mapping $\phi(\cdot)$, and therefore the above objective function is rewritten as:

$$\arg\min : \frac{C}{2}\|\boldsymbol{Y} - \boldsymbol{\Phi}\boldsymbol{w}\|^2 + \frac{1}{2}\|\boldsymbol{w}\|^2 \tag{2}$$

where $\boldsymbol{\Phi}$ is the matrix of mapped samples, $[\phi(\boldsymbol{x}_1), \ldots, \phi(\boldsymbol{x}_n)]$. Applying the Representer's theorem and kernel trick, $\boldsymbol{w} = \boldsymbol{\Phi}^{\mathrm{T}}\boldsymbol{\beta}$ and $\boldsymbol{K} = \boldsymbol{\Phi}\boldsymbol{\Phi}^{\mathrm{T}}$, the solution of the objective function can be obtained by equating its derivative to zero, expressed as:

$$\boldsymbol{\beta} = \left(\boldsymbol{K} + \frac{\boldsymbol{I}}{C}\right)^{-1}\boldsymbol{Y} \tag{3}$$

The overall prediction can be therefore expressed as:

$$\widehat{\boldsymbol{Y}} = \boldsymbol{\Phi}_*\boldsymbol{w} = \boldsymbol{\Phi}_*\boldsymbol{\Phi}^{\mathrm{T}}\boldsymbol{\beta} = \boldsymbol{K}_*\left(\boldsymbol{K} + \frac{\boldsymbol{I}}{C}\right)^{-1}\boldsymbol{Y} \tag{4}$$

where $\boldsymbol{K}_*$ represents the kernel matrix between model input $\boldsymbol{x}_*$ and training samples $\boldsymbol{X}$. Furthermore, Huang et al. (2011) proposed a kernel version of extreme learning machine (KELM), which interprets the KRR in terms of neural network. More details of KRR can be found in Saunders et al. (1998) and (Huang et al. 2011).

In addition, weights can be assigned to different training samples to address the issue of imbalanced dataset and outliers, and therefore the objective function in (2) can be rewritten as:

$$\begin{cases} \arg\min : \frac{1}{2}\|\beta\|^2 + \frac{C}{2}\sum_{i=1}^{n}\lambda_i {e_i}^2 \\ \text{s.t. } e_i = y_i - \widehat{y}_i, i = 1,2,\ldots,n \end{cases} \tag{5}$$

where λ_i represents the weight factor assigned to the ith training sample. Denote the weight matrix as:

$$\lambda = \begin{bmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n \end{bmatrix} \tag{6}$$

The solution to the output weight $\boldsymbol{\beta}$ in (3) can be rewritten as:

$$\boldsymbol{\beta} = \left(\lambda \boldsymbol{K} + \frac{\boldsymbol{I}}{C}\right)^{-1} \lambda \boldsymbol{Y} \tag{7}$$

2.1.2 Multiple Kernel Learning

As stated in Cortes and Vapnik (1995), the performances of kernel-based algorithms are highly depend on the selection of kernel functions. Therefore, MKL is used as an alternative approach to reduce the effort devoted to searching for an optimal kernel (Zhang et al. 2024a, b). It replaces the single kernel function as a linear combination of different base kernels, and thus the overall kernel matrix of MKL takes the form:

$$\boldsymbol{K} = \sum_{p=1}^{m} \gamma_p \boldsymbol{K}_p \tag{8}$$

where γ_p represents the weight assigned to the base kernel $\boldsymbol{K}_p$. The base kernel functions can either be different kernel functions or same kernel function with various Furthermore, according to Liu et al. (2015) and (Bisoi et al. 2020), the L_1 norm and non-negative constraint were imposed on the combination weight γ_p, described as following:

$$\sum_{p=1}^{m} \gamma_p = 1, \gamma_p \geq 0, \forall p \tag{9}$$

In this study, the data-dependent kernel uses a combination of Gaussian functions with various kernel length, given as:

$$K_{\text{Gaussian}}\left(X_i, X_j\right) = e^{-\sigma \|X_i - X_j\|^2} \tag{10}$$

where σ is the length of Gaussian kernel. In this study, 12 different σ values were used, $[0.1, 0.5, 1, 2, 3, \ldots, 10]$, and therefore the overall kernel function of MKL in this study is a linear combination of 12 Gaussian kernels. According to Kloft et al. (2011), when the output vector $\boldsymbol{\beta}$ is given, the weight vector $\boldsymbol{\gamma} = \left\{\gamma_p\right\}_{p=1}^{m}$ can be updated following the equation presented as:

$$\gamma_p = \frac{\boldsymbol{\beta}^{\mathrm{T}} \boldsymbol{K}_p \boldsymbol{\beta}}{\sum_{i=1}^{m} \boldsymbol{\beta}^{\mathrm{T}} \boldsymbol{K}_i \boldsymbol{\beta}} \tag{11}$$

Therefore, the MKL can be obtained by iteratively updating the weight vector $\boldsymbol{\gamma}$ and $\boldsymbol{\beta}$ following Eqs. (7) and (11). The pseudo code describing the training process of W-MKL is presented in Algorithm 1.

Algorithm 1: MKL training

Input: Training data: $\{\boldsymbol{x}_i, y_i\}_{i=1}^{n}$, regularization coefficient: C
Output: $\boldsymbol{\gamma}$ and $\boldsymbol{\beta}$
[1]Construct kernel matrix $\boldsymbol{K} = \sum_{p=1}^{m} \gamma_p \boldsymbol{K}_p$
[2] $q=0$
Repeat
  [3]Update $\boldsymbol{\beta}(q)$ using eq. (7)
  [4]Update $\boldsymbol{\gamma}(q)$ using eq. (11)
  [5] $q=q+1$
Until $\|\boldsymbol{\beta}(q) - \boldsymbol{\beta}(q-1)\|^2 < \varepsilon$
Return $\boldsymbol{\gamma}$ and $\boldsymbol{\beta}$

2.2 Adaptive Weighting

According to the objective function presented in Eq. (5), it is obvious that KRR can better predict those samples with larger weights. Traditional fixed weighting strategy rigidly apply the same set of weights could adversely affect the learning performance. To improve this situation, an adaptive weighting approach is proposed, in which the weights vary adaptively with respect to model input. It can be easily understood that the outputs of those samples with very high similarity were close. Therefore, if we assign larger weights to those samples with higher similarity, the model accuracy could be increased. Moreover, as mentioned earlier in the introduction section, the blasting operations are uncertain due to its nature and this also need to be taken into consideration. Hence, the proposed adaptive weighting strategy considered two concepts when assigning weights to a training sample, namely its correlation to the model input, as well as its uncertainty degree. Specifically, the uncertainty degree of a training sample is measured via a MLOF and the correlation between model input and training sample is evaluated by a weighted Euclidean distance method. The MLOF method is introduced in the first following section, and the overall adaptive weighting strategy is presented in Sect. 2.2.2.

2.2.1 Modified Local Outlier Factor

(a) *Local Outlier Factor:* The local outlier factor (LOF) is a commonly used outlier identification method, which is based on the density of measurements. Given a data point $o\left(\boldsymbol{x}_o, y_o\right) \in D$, the LOF of which can be calculated following the steps presented below:

1. Given k, the k-distance neighborhood of $o(\boldsymbol{x}_o, y_o)$ can be represented by $N_k(o)$, described as:

$$N_k(o) = \{q \in D | d(q,o) \le d_k(o)\} \tag{12}$$

where $d(q,o)$ represents the Euclidean distance between points $q(\boldsymbol{x}_q, y_q)$ and $o(\boldsymbol{x}_o, y_o)$. $d_k(o)$ is the k -distance, defined as the Euclidean distance between the kth nearest neighbor $k(\boldsymbol{x}_k, y_k)$ and point $o(\boldsymbol{x}_o, y_o)$.

2. For a data point $q(\boldsymbol{x}_q, y_q)$ in the neighborhood $N_k(o)$, the reachable distance between o and q is given by:

$$r - \text{dist}(o,q) = \max\{d_k(q), d(o,q)\} \tag{13}$$

Note that the $d_k(q)$ is the k-distance of q, not data point o.

3. Thereafter, the local reachability density of $o(\boldsymbol{x}_o, y_o)$ is calculated as:

$$lrd(o) = \frac{|N_k(o)|}{\sum_{q \in N_k(o)} r - dist(o,q)} \tag{14}$$

where $|N_k(o)|$ represents the number of data points in the neighborhood.

4. Finally, the LOF of the given data point $o(\boldsymbol{x}_o, y_o)$ can be obtained by:

$$lof(o) = \frac{\sum_{q \in N_k(o)} lrd(q)}{|N_k(o)| * lrd(o)} \tag{15}$$

According to these equations, the LOF determines the density level of a data point compared to its neighbors. A larger value of LOF indicates that this data point is obviously less dense than its neighbors and therefore it is more likely to be affected by noise.

(b) *Modified local outlier factor:* For high dimensional data, the traditional Euclidean distance might fail to identify the outliers, since it treats all the attributes of input equally whereas in practical applications they may contribute to the output to different degrees. Therefore, to address this issue, this study fitted a linear equation to measure the dependency between model input and output. For a M-dimensional input vector, the linear equation can be written as:

$$\hat{y} = A \cdot \boldsymbol{x} \tag{16}$$

where $A = \begin{bmatrix} a_1 & \cdots & a_m \end{bmatrix}$ and $\boldsymbol{x} = \begin{bmatrix} x_1 & \cdots & x_m \end{bmatrix}^{\mathrm{T}}$. For two data point $(\boldsymbol{x}_a, y_a)$ and (X_b, y_b), the Euclidean distance between them was modified as:

$$d(a,b) = \sqrt{(\hat{y}_a - \hat{y}_b)^2 + (y_a - y_b)^2} \tag{17}$$

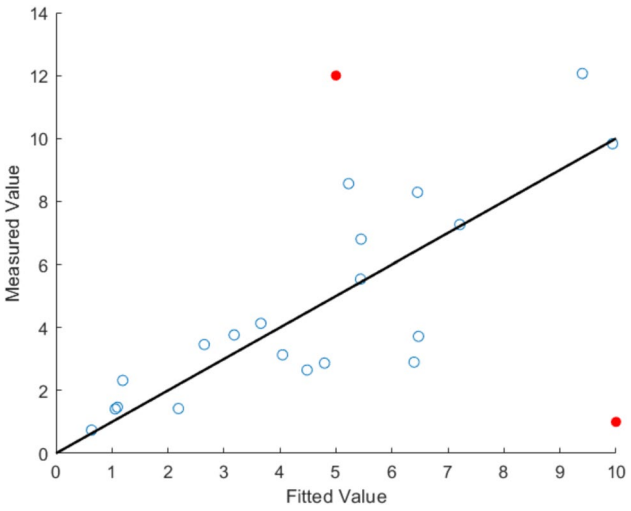

Fig. 1 Fitted value versus measured value of linear equation

The original model input X is replaced by the fitted value $\hat{y}$. Using the modified Euclidean distance presented in Eq. (17) in the calculation of Eqs. (12) and (13), the MLOF can therefore be obtained following the step (1)–(4) illustrated in the previous section.

Using the fitted linear equation, one can map the original high dimensional data $(\boldsymbol{x}, y)$ into two-dimensional space, denote as $(\hat{y}, y)$. Therefore, the proposed MLOF method can be explained graphically in two-dimensional space. Figure 1 demonstrated an example of some data points and their fitted equation.

It is acknowledged that the datapoints should be scattered along the fitted line since they follow the same relationship, whereas those datapoints located far from the fitted line were considered to have more uncertainty. Therefore, by mapping the original data point (X, y) into $(\hat{y}, y)$, those samples that defy the common sense can be easily identified, as the two red dots shown in Fig. 1. It is worth mentioning that the MLOF method is not used to determine whether a datapoint is an outlier, but rather to measure the extent to which this datapoint obeys the general relationship between model output and inputs, as such kind of datapoint can reduce the prediction accuracy and should be assigned with smaller weights.

2.2.2 Adaptive Weighting

Given a model input $\boldsymbol{x}_t$, the weight matrix $\boldsymbol{\lambda}$ can be obtained based on the calculation of similarity vector S:

$$S = \begin{bmatrix} s_{t,1} \\ s_{t,2} \\ \vdots \\ s_{t,n} \end{bmatrix} \tag{18}$$

where $s_{t,i}$ represents the similarity between model input $\boldsymbol{x}_t$ and training data $\boldsymbol{x}_i$. The similarity is calculated based on the weighted Euclidean distance and aforementioned MLOF.

(a) *Weighted Euclidean distance:* Usually, one can measure the similarity between data using Euclidean distance. As stated previously, the traditional Euclidean distance considered every dimension of the sample has equal importance and therefore is not suitable. To address this issue, a weighted Euclidean distance is used to measure the similarity between data. The weighted Euclidean distance can be calculated as:

$$d_{i,j} = \sqrt{\widetilde{r_1}\left(x_{t,1} - x_{i,1}\right)^2 + \widetilde{r_2}\left(x_{t,2} - x_{i,2}\right)^2 + \cdots + \widetilde{r_m}\left(x_{t,m} - x_{i,m}\right)^2} \tag{19}$$

In the above equation, $\widetilde{r_{\mathrm{m}}}$ represents the influence degree of mth attribute. To determine $\widetilde{r_{\mathrm{m}}}$, the Pearson correlation coefficient is used, and therefore $\widetilde{r_{\mathrm{m}}}$ is the normalized Pearson correlation coefficient, calculated by:

$$\widetilde{r_{\mathrm{m}}} = M * \frac{|r_m|}{\sum_{m=1}^{M} |r_m|} \tag{20}$$

(b) *Adaptive weighting based on weighted Euclidean distance and MLOF:* According to Eq. (19), a training sample with smaller the weighted Euclidean distance value represents a higher correlation to the model input and therefore should be assigned with larger weight to guarantee for a better fitting performance. Moreover, the level of uncertainty of this sample also needs to be taken into consideration to avoid the aforementioned negative impact, and therefore its MLOF value should be referred as well. Taking into account of these two aspects, Eq. (21) is therefore introduced to calculate the general similarity factor between $\boldsymbol{x}_t$ and $\boldsymbol{x}_i$:

$$s_{t,i} = e^{-d_{t,i}{}^2 * lof_i} \tag{21}$$

where lof_i is the MLOF value calculated for the training sample $\boldsymbol{x}_i$. With the use of the exponential function, the similarity factor can be scaled into (0, 1]. A training sample with smaller weighted Euclidean distance and MLOF value is considered to be more reliable and therefore has a higher similarity value. On the other hand, if a training sample has higher correlation to model input but located far from its nearby samples, its general similarity factor would not be high. The weight factor $\lambda_{t,i}$ is the normalized similarity factor, defined as:

$$\lambda_{t,i} = n \frac{s_{t,i}}{\sum_{i=1}^{n} s_{t,i}} \tag{22}$$

2.3 Adaptive Weighted Multi-kernel Learning

The proposed AW-MKL is a combination of the previously introduced W-MKL, adaptive weighting strategy. During the development of AW-MKL, the weight matrix $\boldsymbol{\lambda}$ in Eq. (6) is first calculated following the adaptive weighting method presented in Sect. 2.2. Subsequently, the MKL model is then trained by iteratively updating the kernel combination coefficient $\boldsymbol{\gamma}$ and output weight $\boldsymbol{\beta}$ based on Algorithm 1. Given a model input $\boldsymbol{x}_t$, the output of AW-MKL can be written as:

$$\widehat{y}_t = \left(\sum_{p=1}^{m} \gamma_p \boldsymbol{K}_{p,*}\right)\boldsymbol{\beta} = \sum_{p=1}^{m} \gamma_p \boldsymbol{K}_{p,*}\left(\boldsymbol{\lambda}\sum_{p=1}^{m} \gamma_p \boldsymbol{K}_p + \frac{\boldsymbol{I}}{C}\right)^{-1} \boldsymbol{\lambda Y} \tag{23}$$

where $\boldsymbol{K}_{p,*}$ is the kernel matrix between model input $\boldsymbol{x}_t$ and all training samples. For a set of testing sample denoted as $\{\boldsymbol{x}_t\}_{t=1}^{L}$, the corresponding outputs of AW-MKL can be obtained following the pseudo code presented in Algorithm 2.

Algorithm 2: AW-MKL

Input: Testing data: $\{\boldsymbol{x}_t\}_{t=1}^{L}$, Training data: $\{\boldsymbol{x}_i, y_i\}_{i=1}^{n}$, regularization coefficient: C, k nearest neighbor: k

Output: Predicted value: $\widehat{\boldsymbol{Y}} = \{\hat{y}_t\}_{t=1}^{L}$

Fit the linear equation in eq. (16) based on $\{\boldsymbol{x}_i, y_i\}_{i=1}^{n}$

Obtain the pair $\{\hat{y}_i, y_i\}_{i=1}^{n}$

for $\boldsymbol{i = 1, \ldots, n}$

Identify $N_k(o)$ in eq. (12) of mapped data point $(\hat{y}_i, y_i)$

for $\boldsymbol{q = 1, \ldots, k}$

Calculate $r - dist(o, q)$ in eq. (13) between $(\hat{y}_q, y_q)$ and $(\hat{y}_o, y_o)$

endfor

Calculate $lrd(o)$ in eq. (14) of $(\hat{y}_o, y_o)$

endfor

Calculate the MLOF values $\{lof(o)\}_{o=1}^{n}$ using eq. (15)

for $\boldsymbol{t = 1, \ldots, L}$

for $\boldsymbol{i = 1, \ldots, n}$

Calculate weighted Euclidean distance $d_{t,\mathrm{i}}$ between $\boldsymbol{x}_t$ and $\boldsymbol{x}_i$ using eq. (19)

Calculate similarity factor $s_{t,i} = e^{-{d_{t,i}}^2 * lof_i}$

Add $s_{t,i}$ to the similarity matrix S

endfor

Calculate the weight matrix $\boldsymbol{\lambda}$ based on S and eq. (22)

Call **Algorithm 1** to obtain $\boldsymbol{\gamma}$ and $\boldsymbol{\beta}$

Calculate the output $\hat{y}_t$ using eq. (23)

endfor

Return $\widehat{\boldsymbol{Y}} = \{\hat{y}_t\}_{t=1}^{L}$

3 Case Study and Model Development

3.1 Data Collection

The data used in this study is collected from Sungun Copper Mine site, which is one of the largest porphyry copper mines in Iran. It is located in the East Azarbaijan province, at 46′43°E longitude and 38′42°N latitude, and is 2000 m above sea level. Copper is the main product of this mine, which can be extracted from its primary minerals, i.e., chalcopyrite, pyrite, chalcocite, cuprite and malachite. In addition to copper, gold, silver, and molybdenite can also be extracted from this mine. Table 1 provides a summary of the geologic and geometrical features of this mine.

In Sungun Copper Mine, bench blasting with one free face is used for mineral extraction process. Main explosive material used was ANFO, and the blastholes were stemmed with drill-cutting particles. Detonating cord was used as initiation of blasting operations. The main blasting pattern was triangular, and the burden to spacing ratio was designed according to the characteristics of the blasting block in different area of the mine. The blasting operations used a flat face method as blasting sequence, where the inter-hole delay time is set as zero and a delay happens between blasting

rows. Flyrock is considered as one of the main ill effects of the blasting operations in this mine. As a result, this study attempted to develop a model predicting flyrock distance. For this purpose, 234 blasting events were recorded for model development.

3.2 Descriptive Analysis

The dataset used in this study consists of 6 blasting parameters obtained from these blasting events, i.e., hole length, spacing, burden, stemming, specific charge, and specific drilling. The basic descriptive statistics of these features as well as the measured flyrock distance were summarized in Table 2. Note that the measured flyrock distance refers to the maximum flyrock distance recorded in a single trial blast. Figure 4 illustrated the correlation matrix plot of the dataset, which can be used to study the level of relationship between target (flyrock distance) and its six influencing parameters.

According to the correlation coefficients demonstrated in Fig. 2, other than specific charge and specific drilling, the other features have an inverse relationship with flyrock distance. Furthermore, it can be observed that specific charge shows the most significant influence on flyrock distance. Compared with other features, specific drilling has the lowest correlation to flyrock distance, whose correlation coefficient is 0.1009, indicating a less significance, and therefore is not considered as the input of model. As a result, the models developed in the following section only use hole length, spacing, burden, stemming and specific charge as model input to estimate flyrock distance.

3.3 Evaluation of Model

In this study the performances of the models are evaluated by three evaluation indices, root mean squared error (RMSE), mean absolute error (MAE) and variance accounted for (VAF). The formulas used to calculate these indices are given as follows:

$$\mathrm{RMSE} = \sqrt{\frac{\sum_1^n \left(\widehat{y}_i - y_i\right)^2}{n}} \tag{24}$$

$$\mathrm{MAE} = \frac{\sum_{i=1}^n \left|\widehat{y}_i - y_i\right|}{n} \tag{25}$$

$$\mathrm{VAF} = \frac{1 - \mathrm{var}(y_i - \widehat{y}_i)}{\mathrm{var}(y_i)} \times 100 \tag{26}$$

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \widehat{y}_i)^2}{\sum_{i=1}^n (y_i - \overline{y})^2} \tag{27}$$

where $\widehat{y}_i$ and y_i represent the predicted and observed value respectively, and $\overline{y}$ stands for the average value of y_i. The value of n represents the number of testing samples.

3.4 Model Development

The blasting dataset collected were randomly divided into training and testing datasets for model development and evaluation purpose. The training dataset consists of 200 sets of data, and the remaining 34 sets of data were used to examine the performance of the models. In this section, the development of the proposed AW-MKL model is first illustrated, followed by the implementation of several comparative methods including KRR, SVM, RF, GBDT, M5 tree and MARS. Since the performances of the aforementioned models are affected by their hyperparameters, the fivefold cross validation method and grid search were used for hyperparameters tuning purpose. The validation dataset consisted of 20 sets of data selected randomly from training dataset, that is, during the cross-validation process, 180 sets of data served as the training dataset and 20 sets of data were used for validation purpose. The optimal hyperparameter values were determined by mean validation RMSE.

As mentioned previously, the proposed AW-MKL used a MKL method to reduce the effort devoted in tuning an optimal kernel. Therefore, only two parameters need to be determined during this stage, one is the regularization coefficient C and another one is the value of k-nearest neighbor included in the calculation of MLOF. The value of k was determined in the first place. To maximize the distinction between the MLOF value, the k value that achieved the maximum interquartile range (IQR) was selected. The IQR represents the difference between the third and first quartiles, defined as:

$$IQR = q_3 - q_1 \tag{28}$$

where q_3 and q_1 are the third and first quartiles respectively. The IQR values under k ranging from 4 to 10 were presented in Fig. 2.

Therefore, the value of k-nearest neighbor was determined as 4 based on the results presented in Fig. 3. Subsequently, the aforementioned fivefold cross validation process was implemented to obtain the optimal regularization coefficient. Considering the characteristic of C, it was searched in the range $[1, 50, 100, 150, 200, \ldots, 500]$. According the validation results, the optimal value of was determined as 500.

To demonstrate the superiority of the proposed method, several existing machine learning algorithms were implemented as comparison, including KRR, SVM, RF, GBDT, M5 tree, MARS and AdaBoost. Among them, KRR and SVM are also non-parametric models using the kernel trick. RF, GBDT, M5 tree and AdaBoost were tree-based models. RF, GBDT, and AdaBoost are ensemble learning algorithms

Table 1 Geologic and geometrical features of Sungun Copper Mine (Yari et al. 2023)

Geologic/geometrical properties	Value
Geologic reserve of the deposit	796MT
Proved reserve	410MT
Average grade	0.67%
Height of the working benches	12.5 m
Slope of the working benches	68′
Angle of the overall pit slope	37′
Width of the ramp	30 m
Slope of the ramp	5′
Age of the mine	Around 32 yr
Overall stripping ratio	1.7

Table 2 Description of dataset

	Unit	Min	Max	Mean	Standard deviation
Hole length	m	10	14	12.31	1.18
Spacing	m	2	6.5	4.53	0.90
Burden	m	2	5	3.69	0.82
Stemming	m	1.8	4.5	3.66	0.76
Specific charge	kg/mm^3	0.2	0.93	0.46	0.20
Specific drilling	m	0.04	0.29	0.07	0.04
Flyrock distance	m	10	100	68.68	17.42

based on bagging and boosting methods, while the M5 tree is a model tree algorithm. MARS is also a non-parametric regression technique which can be seen as an extension of linear models that simulate the non-linearity and interaction between variables. The developments of these models were demonstrated as following and the obtained optimal hyperparameters were summarized in Table 3. A flowchart was presented in Fig. 4 illustrating the steps involved in this study.

1. *KRR:* The KRR developed in this study used a single Gaussian function as its kernel function. Based on the previous description, hyperparameters affecting KRR are regularization coefficient C and kernel length σ. Similar to the proposed AW-MKL, the searching space of C was $[1, 50, 100, 150, 200, \ldots, 500]$. The value of σ was determined in the range $[0.1, 0.5, 1, 2, 3, \ldots, 10]$. The cross-validation process reported the optimal pair $\{C, \sigma\}$ was $\{500, 6\}$.
2. *SVM:* As mentioned earlier, SVM is also a kernel-based machine learning method and the kernel function of which was again selected as Gaussian function. The SVM is similar to KRR in some ways, since their objective functions are close. Both of them use a regularization coefficient C to control the trade off between structural risk and training error. Moreover, SVM introduces an ε-insensitive band to allow for an over/underestimation of training sample within $\pm\varepsilon$. There-

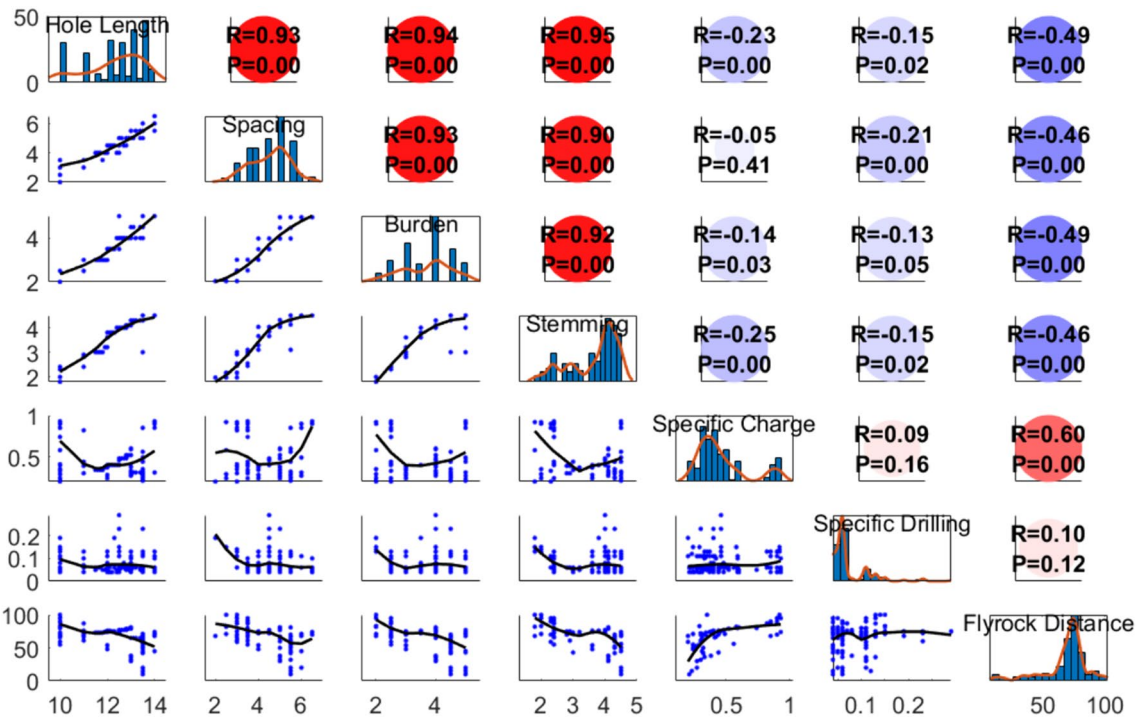

Fig. 2 Correlation matrix of dataset

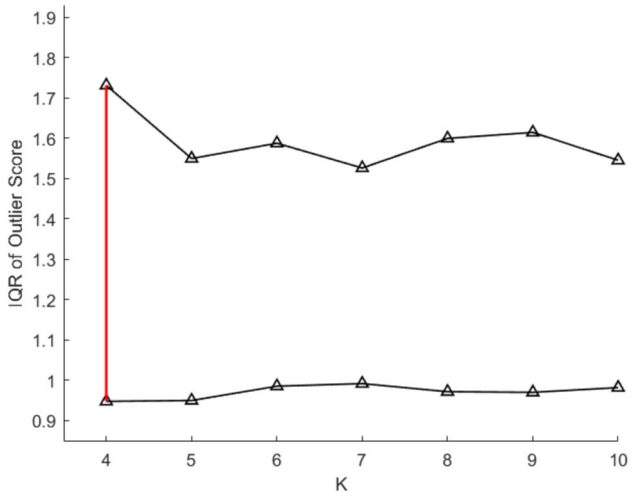

Fig. 3 IQR of modified outlier factor under various k

Table 3 Optimal hyperparameters of developed models

	Hyperparameters
AW-MKL	$C = 500, k = 4$
KRR	$C = 500, \sigma = 6$
SVM	$C = 50, \sigma = 0.5, \varepsilon = 2^{-7}$
RF	ntree = 200, mtry = 4
GBDT	ntree = 450, $\alpha = 0.6$
MARS	nprune = 15, degree = 3
M5 tree	depth = 10, minleafsize = 10
AdaBoost	$nrf = 100, \alpha = 1$

fore, other than C and σ, the half bandwidth ε of SVM also needs to be tuned. Considering the ε should be a very small constant, and it is therefore determined from $\left[2^{-10}, 2^{-9}, 2^{-8}, \ldots, 2^{-1}\right]$. Finally, the optimal values of C, σ and ε were determined as 50, 0.5 and 2^{-7}.

3. *RF:* The RF is an ensemble learning algorithm based on the concept of bagging. It trains a given number (ntree) of decision trees based on ntree sub-training datasets produced using bootstrap sampling method. To further increase the variance of the developed trees, RF limits the number of attributes selected in the split of parent node, that is, in each split of parent node, RF randomly select a number (mtry) of attributes from all attributes to search for the best split. Hence, the number of trees (ntree) and the number of attributes selected in each split (mtry) were tuned to obtain an optimal RF. ntree was searched in the range $[10, 50, 100, 150, \ldots, 500]$. Since only five features were used as model input, the mtry was selected from 2 to 5. The validation results reported that a combination of $\{\text{ntree} = 200, \text{mtry} = 4\}$ was the optimum.
4. *GBDT:* The GBDT is also an ensemble learning algorithm based on decision tree. Unlike RF, GBDT is a collection of ntree decision trees developed using L_2 -Boosting. In GBDT, decision tree is trained iteratively to compensate for the residuals of the previous model and a learning rate α is added used to control the convergence speed of GBDT. Therefore, the adjustable hyperparameters of GBDT are the number of trees (ntree) and learning rate (α). The searching space of ntree was set

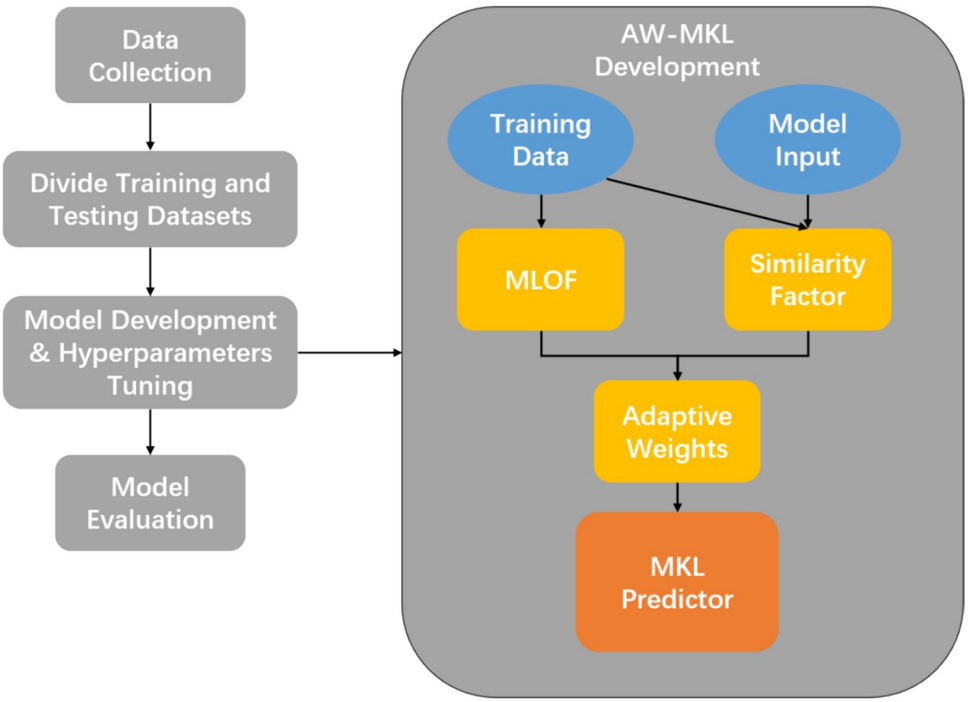

Fig. 4 Flowchart of this study

Table 4 Evaluation metrics obtained for all models

	Training performance								
	RMSE	Rank	MAE	Rank	VAF	Rank	R^2	Rank	Overall rank
AW-MKL	1.50	2	0.57	2	99.96	2	0.9923	2	2
SVM	2.13	4	1.05	4	99.91	4	0.9855	5	4
KRR	1.80	3	0.86	3	99.93	3	0.9884	3	3
RF	3.75	6	2.11	5	99.72	6	0.9606	6	6
GBDT	0.43	1	0.09	1	99.99	1	0.9994	1	1
MARS	5.05	7	3.41	7	99.49	7	0.9181	7	7
M5 tree	6.57	8	3.82	8	99.14	8	0.8612	8	8
AdaBoost	2.46	5	2.24	6	99.88	5	0.9857	4	5
	Testing performance								
	RMSE	Rank	MAE	Rank	VAF	Rank	R^2	Rank	Overall rank
AW-MKL	2.05	1	0.98	1	99.92	1	0.9827	1	1
SVM	2.84	2	1.54	2	99.84	2	0.9706	2	2
KRR	3.13	3	1.70	3	99.80	3	0.9606	3	3
RF	7.59	7	3.46	6	98.84	7	0.7767	7	7
GBDT	6.02	6	1.81	4	99.27	6	0.8588	6	5
MARS	5.90	5	4.01	8	99.30	5	0.8616	5	6
M5 tree	8.65	8	3.54	7	98.50	8	0.7128	8	8
AdaBoost	5.08	4	2.98	5	99.48	4	0.8969	4	4

as same as RF. As for α, according to Caponnetto and De Vito (2007) and Lin et al. (2017), $\frac{1}{2} \leq \alpha \leq 1$ is recommended, and therefore the searching space of α was $[0.5, 0.6, 0.7, \ldots, 1]$. According to the validation results, the optimal ntree and α were 450 and 0.6, respectively.

5. *MARS:* MARS is a multivariate nonparametric predictive technique developed on a series of piecewise local linear model or cubic segments. A two-step training procedure is used in MARS, where basis functions are added iteratively to obtain the lowest training error in the first stage and those least contributing basis functions

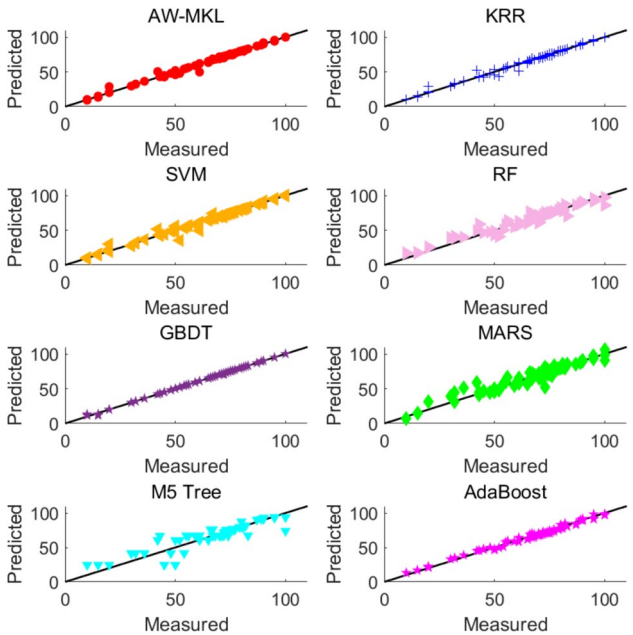

Fig. 5 Training performances of models

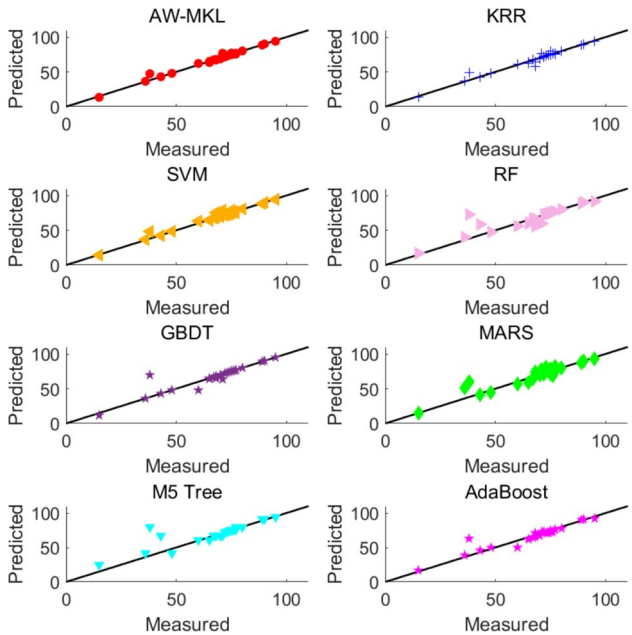

Fig. 6 Testing performances of models

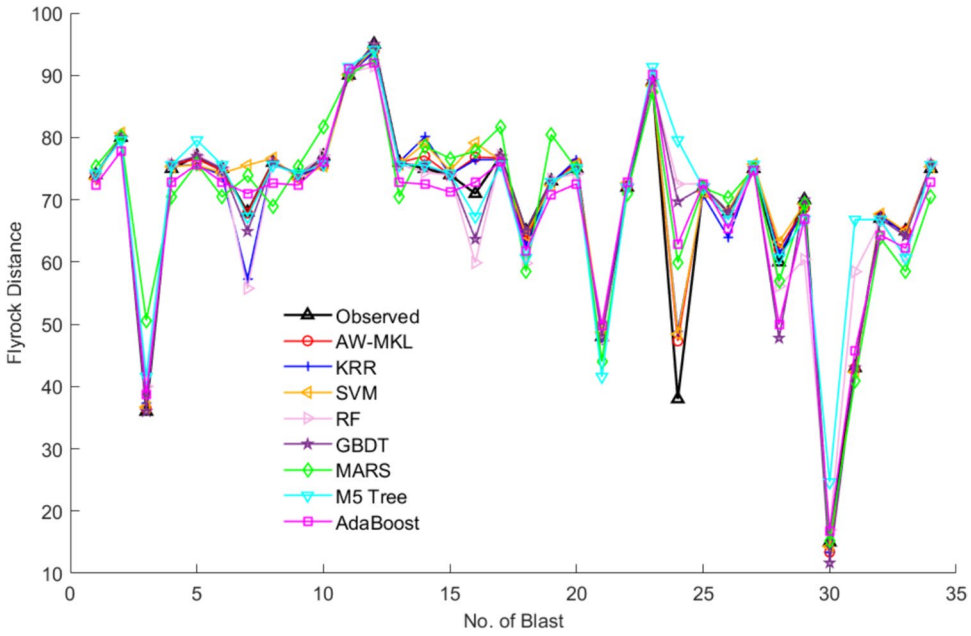

Fig. 7 Predicted versus measured testing results

are removed in the subsequent prune stage. Furthermore, MARS also allows for the interaction between different features. Therefore, its performance is affected by the maximum degree of interaction (degree) and maximum number of terms after prune (*nprune*). Since the input data involves five features, the value of degree was selected from 1 to 5. Considering the size of dataset, too much basis functions would result in overfitting, and therefore the value of *nprune* was validated in the range $\{2, 5, 10, 15 \ldots, 50\}$. The optimal *nprune* and degree reported by the validation results were 15 and 3 respectively.

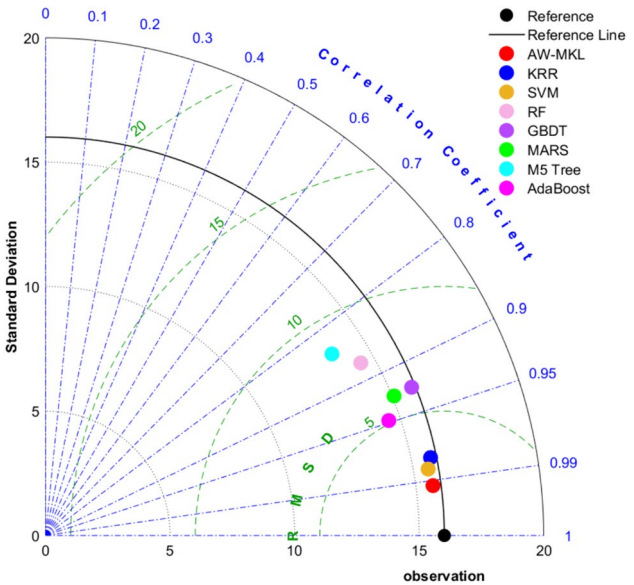

Fig. 8 Taylor plot of testing results

6. *M5 Tree:* Different from RF and GBDT, M5 Tree only consists of a single tree. M5 Tree is a model tree, where the solution of each node is a linear model. Therefore, compared with traditional classification and regression tree (CART) which uses the mean value to represents the samples within the node, M5 Tree has stronger approximation ability. The performance of M5 Tree can be optimized by specifying the complexity of tree. Therefore, in this study, the authors selected two hyperparameters to control the complexity of M5 Tree, which are maximum tree depth (depth) and minimum leaf size (minleafsize). Considering the size of dataset, the searching space was determined as $[2, 3, 4, \ldots, 20]$ and $[5, 10, 20, 30, \ldots, 100]$. After the validation process, the optimal values of these two hyperparameters were determined as 10 and 10 respectively.
7. *AdaBoost:* The AdaBoost model developed in this study follows the same hyperparameters as the one developed in Yari et al. (2023), since the dataset used was identical. The base learner of AdaBoost was selected as RF, and the hyperparameters of RF learner were the same as aforementioned. The number of estimators (*nrf*) was 100 and the learning rate (α) was 1.

4 Results and Discussion

In the previous section, seven machine learning models were developed. To evaluate the accuracy of these models, three evaluation metrics presented in Sect. 3.3 were used. Table 4 summarizes and ranks the evaluation metrics obtained for the developed models. The ranking system assesses the

Table 5 Effect of adaptive weighting

	RMSE (tr)	MAE (tr)	VAF (tr)	RMSE (ts)	MAE (ts)	VAF (ts)
AW-MKL	1.50	0.57	99.96	2.05	0.98	99.92
MKL	1.64	0.71	99.95	2.32	1.24	99.89
MKL (MLOF only)	4.29	1.43	99.63	5.74	2.64	99.34
MKL (weighted Euclidean distance only)	1.52	0.59	99.95	2.11	1.06	99.91
KRR	1.80	0.86	99.93	3.13	1.70	99.80
AW-KRR	1.64	0.71	99.94	2.57	1.39	99.87

performance of models based on their respective standings under each evaluation metric. The final overall ranking is determined by the summation of the model's standings across the three metrics, with the lowest total sum receiving the highest rank. For instance, if a model is ranked 1, 2, 1 under the three metrics respectively, its total sum would be 4. Conversely, another model with rankings of 2, 1, 2 would have a total sum of 5. Therefore, the model with a total sum of 4 would have a higher overall ranking. The predicted versus measure flyrock distances in training and testing datasets were depicted in Figs. 5 and 6. Figure 7 presents the exact fitting performances of models at each testing sample. A Taylor plot is demonstrated in Fig. 8 to further illustrates the performance difference between models on testing dataset,

According to the results presented in Table 4, all the developed models were capable for predicting blast-induced flyrock distance. Among them, the proposed AW-MKL model outperformed other models at testing phase, achieving the highest testing accuracy, with RMSE of 2.05, MAE of 0.98, VAF of 99.92 and R^2 of 0.9827. In training phase, the proposed method only underperformed GBDT, achieving RMSE of 1.50, MAE of 0.57, VAF of 99.96 and R^2 of 0.9923, also demonstrating a high predictive capability. Furthermore, considering the unsatisfactory testing results of GBDT, which only ranked the 4th among all the developed models, the proposed model is therefore emerged as the most reliable model in predicting blast-induced flyrock distance.

Among all the models developed, the performances of RF, MARS and M5 Tree were the worst at both training and testing phase. Based on the Figs. 5 and 6, all of them cannot guarantee for accurate predictions, since their predicted values were scattered along the 45° line. RF and MARS performed better at training dataset as their prediction errors were less obvious in Fig. 5, whereas several datapoints were obviously over/underestimated by M5 Tree. At testing phase, MARS were more reliable than RF and M5 Tree, as the R^2 of those two models were only 0.7767 and 0.7128, indicating less correlation to the ground truth value. According to Fig. 6, it is obvious that the maximum prediction error of MARS was smaller. Furthermore, in the Taylor graph presented in Fig. 8, the point representing MARS were closer to the reference point, which implies the predictions of MARS were more correlated to the measured flyrock distance. Therefore, taking into account of the above discussion, the MARS model is regarded to be superior to the other two models. On the contrary, M5 Tree is the least suitable model to predict the distance of flyrock induced by blasting operation. Moreover, comparing RF with AdaBoost, it is obvious that the predictability of RF is enhanced with the use of adaptive boosting method, as the accuracy at both training and testing phases were significantly improved.

Other than RF, MARS and M5 Tree, all the remaining models demonstrated high predictive capability. As mentioned earlier, the GBDT obtained the best training performance, with RMSE of 0.43, MAE of 0.09, VAF of 99.99 and R^2 of 0.9994, representing a very high accuracy. This can also be observed in Fig. 5, where the predictions of GBDT are almost match the 45° line exactly. However, at testing dataset, GBDT significantly over/underestimated two datapoints according to Fig. 6. For better illustration, the predicted versus measured value at each datapoint was depicted in Fig. 7. Based on this figure, we can see that the GBDT failed to predict No. 24 and 28 testing datapoints, while the remaining models can give closer predictions. At the other datapoints, all of them showed low error rate. Hence, although the developed GBDT model outperformed the proposed method at training phase, it has a high risk of overfitting, whereas the proposed model is able to perform well at both training and testing phase, and is therefore considered more reliable than GBDT. As for KRR and SVM, their overall performances were very close. KRR performed better at training stage, while the SVM showed less error rate at testing dataset. Thus, it is hard to tell whether KRR or SVM is better. Furthermore, compared with other models, the testing performances of kernel-based models (AW-MKL, KRR and SVM) were significantly better. According to the Taylor graph presented in Fig. 8, the performance difference between them was obvious. Since all the kernel-based models performed well, we can also draw a conclusion that the kernel-based methods were more suitable in predicting blast-induced flyrock distance.

Table 6 Error analysis of models

	Max error (e_{max})	Average error ($\bar{e}$)	Uncertainty bandwidth
AW-MKL	9.3177	0.3457	± 4.0294
KRR	10.9217	− 0.0704	± 6.2303
SVM	10.4391	0.8521	± 5.3916
RF	34.9905	0.1792	± 15.0966
GBDT	31.6993	0.1791	± 11.9646
MARS	21.9378	0.5687	± 11.6826
M5 Tree	41.5556	1.8276	± 16.8183
AdaBoost	24.8347	− 0.4982	± 10.0602

To illustrate the effect of the proposed adaptive weighting method, the several MKL models were developed for comparison purpose. One was a single MKL model without any weighting strategy, other one was developed using the MLOF for the similarity calculation in Eq. (21) only, and another one used the weighted Euclidean distance only. Moreover, the authors also developed an AW-KRR model. All the hyperparameters of the aforementioned models were determined as the same as the AW-MKL or KRR model, that is, $C = 500$, $k = 4$ and $\sigma = 0.6$. The results of them were presented in Table 5.

Comparing the performances between KRR, MKL and AW-KRR, it is obvious that both the adaptive weighting strategy and MKL are able to improve the performance of KRR. Therefore, the combination of adaptive weighting and MKL method can maximally enhance the performance of KRR. Since the proposed adaptive weighting strategy consists of two parts, weighted Euclidean distance and MLOF, the effect of them were compared respectively. According to Table 5, it would not help enhancing the performance of MKL if one only uses the MLOF to generate sample weights, adversely, the model performances would become even worse. On the contrary, generating sample weights based on weighted Euclidean distance is feasible. In fact, the improvement made by the weighted Euclidean distance weighting is of the most significance, which illustrated that the weighted Euclidean distance can indeed reflect the correlation between samples. After added with MLOF, the model accuracy was slightly improved, which implies that the combination of MLOF and weighted Euclidean distance is feasible. This is probably because some datapoints with relatively high similarity might violate the general relationship between AOp and its influencing factors and therefore are not helpful in estimating the unseen data. Hence, adding the weighted Euclidean distance based weighting method with MLOF can reduce the impact brought by those samples that are less correlated to other datapoints. As a result, based on the previous discussion, it is obvious that the proposed adaptive weighting strategy can effectively improve the performance of MKL model.

To measure the uncertainty of these models in predicting blast-induced flyrock distance, the error analysis of models was carried out and the results of which were summarized in Table 6, including the maximum error value, average error and the uncertainty bandwidth. The uncertainty band refers to the 95% error confidence band (Newcombe 1998), calculated as:

$$\text{uncertainty bandwidth} = \pm 1.96 S_e \tag{29}$$

where S_e represents the standard deviation of error, defined as:

$$S_e = \sqrt{\frac{\sum_{i=1}^{n} (e_i - \bar{e})^2}{n-1}} \tag{30}$$

$$\bar{e} = \frac{\sum_{i=1}^{n} e_i}{n} \tag{31}$$

In the above equations, e_i is the error of the ith testing sample, which is defined as $e_i = y_{\text{predicted}} - y_{\text{observed}}$.

Based on the results presented in Table 6, other than KRR model, all the other models obtained a positive value of $\bar{e}$, indicating an overestimation of these models. The maximum error value of AW-MKL is the lowest (9.3177), whereas the M5 Tree significantly overestimate the induced flyrock distance by 41.556 m. Meanwhile, it is also obvious that the maximum error values of the kernel-based models, i.e., AW-MKL, SVM and KRR, were all around 10, which were significantly lower than the other models. Beyond that, the proposed AW-MKL model achieved the lowest uncertainty bandwidth, with ± 4.0294, which means that the proposed model is the most reliable among all the developed models. On the contrary, the predictions of RF and M5 Tree models are more uncertain as their uncertainty band were relatively wide (± 15.0966 and ± 16.8183 respectively).

Furthermore, to examine the significance of the input variables, a sensitivity analysis was performed based on Morris method (Morris 1991) and the proposed AW-MKL model. The analysis results are presented in Fig. 9.

According to the results of Morris method, other than specific charges, the other features exhibit a negative relationship to the flyrock distance. Among them, the importance level of spacing and stemming are the most significant. Moreover, specific charge is the only feature with positive importance value, which means the flyrock distance increases when the value of specific charge increases. Meanwhile, it also has the maximum absolute importance level, indicating that it has the most significant impact on the blast-induced flyrock distance. Furthermore, comparing the Morris results with the previous correlation analysis presented in Sect. 3.2, similar conclusion can be drawn. Both of them indicate that specific charge has the biggest influence on

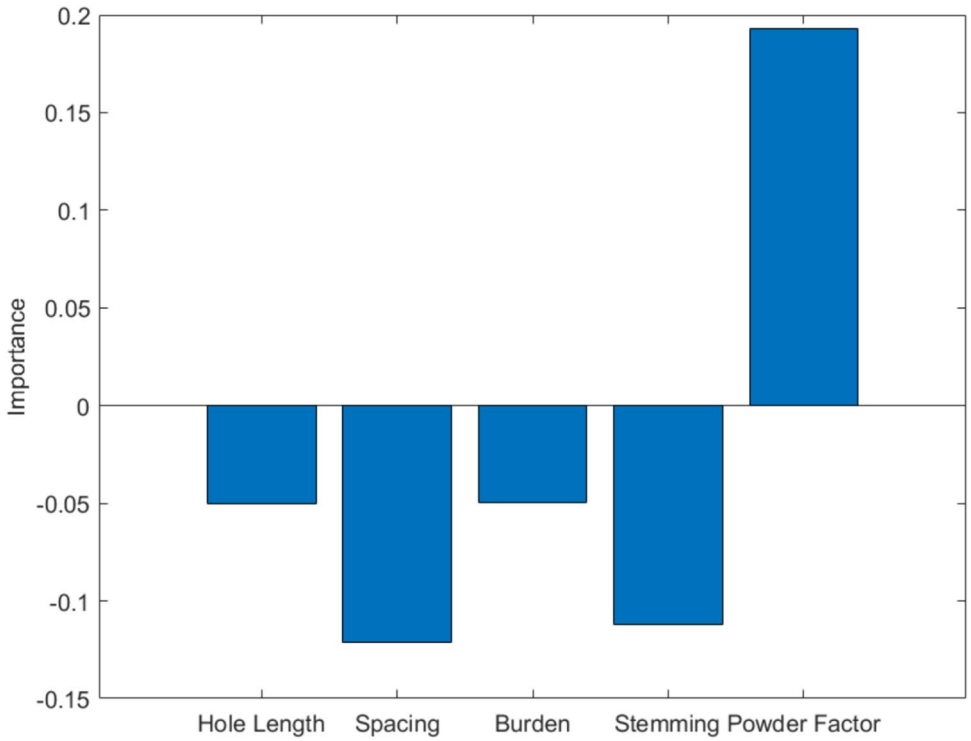

Fig. 9 Sensitivity analysis of input features

flyrock distance and exhibit a direct relationship, while the remaining features are inversely related to the output change.

In summary, all the developed models showed good potential for predicting the blast-induced flyrock distance. Among all the developed models, the proposed AW-MKL was superior to other models, for its strong approximation ability and high reliability. Moreover, the proposed adaptive weighting strategy which uses the combination of sample similarity weighting and MLOF was also proved to be effective in improving the performance of MKL model.

5 Conclusion

In this study, an attempt has been made to predict blast-induced flyrock distance. For this purpose, a total number of 234 sets of blasting data were collected from Sungun Copper Mine site to develop the predictive model. Considering the size of dataset as well as the uncertainty involved in blasting operation, an AW-MKL model was proposed to achieve higher prediction accuracy and reliability. The proposed method uses the combination of adaptive weighting strategy and MKL method. The weighting strategy is able to reallocate the model attention by taking into account of the correlation between input and training samples as well as the degree that the sample obey the general distribution. To evaluate the effectiveness of the proposed approach, six existing machine learning models were also implemented as comparison, i.e., KRR, SVM, RF, GBDT, MARS and M5 Tree. For model development and evaluation purpose, 200 sets of blasting data were randomly selected as training dataset, and the remaining 34 sets of data were used for the testing purpose. During the modeling procedure, a fivefold cross validation was used with grid search method to determine the optimal hyperparameters of the aforementioned models, where each fold of validation dataset consists of 20 sets of data randomly selected from training datasets. After the optimal hyperparameters were obtained, these models were trained again on the whole training dataset and subsequently evaluated on the testing dataset according to three evaluation indices, RMSE, MAE and VAF.

According to the evaluation and discussion of these models, the proposed AW-MKL was considered as the most superior model in predicting blast-induced flyrock distance in Sungun Copper Mine site, which achieve the second-best training performance and highest testing accuracy with RMSE of 1.50/2.05, MAE of 0.57/0.98 and VAF of 99.96/99.92, demonstrating very high reliability and estimation ability. Although GBDT performed better at training stage, its prediction on unseen data was less reliable due to its high uncertainty. Furthermore, by comparing the performance between AW-MKL, KRR, AW-KRR and several MKL models, both MKL method and the proposed weighting strategy are proved to be capable for improving the performance of KRR. Meanwhile, the comparison results also demonstrate the theory behind the proposed adaptive weighting method, where the weighted Euclidean distance can reflect the correlation between samples and MLOF is

capable for weaken the impact of those samples with less uncertainties.

The main difficulty encountered in the development of the proposed model is the problem of small dataset. Due to funding, environmental and regulation issues, mining companies would not conduct thousands of trial blasts Hence, the sizes of blasting datasets are normally small. As a result, the main challenge in this field of study is to improve the prediction performance on small dataset. In this study, the adaptive weighting method is proposed to address this issue. Although it demonstrated satisfactory prediction accuracy, there are still some drawbacks and limitations that need to be improved in future. First, the used dataset only consists of blasting data collected from one site, and therefore the predictive capability of the developed model has not been evaluated in other sites. Apart from that, in the collected dataset, only the blasting parameters are available, whereas the uncontrollable parameters describing the variation of geologic condition such as rock quality designation (RQD) are not included. Since the geologic condition also affects the result of a blasting operation, and this would limit the accuracy and the applicability of the model in other area with different geologic conditions. As a result, future study can focus on extending the applicability of the developed model by integrating various blasting datasets. The combined dataset consists of blasting data collected from various sites with different geologic conditions, and thus the developed model can be applicable under various scenarios. This also allows for developing deep learning models or the application of transfer learning methods. Moreover, analyzing the causation of flyrock phenomenon is also of importance. Identifying the causative factors of flyrock phenomenon and quantifying their relationship with the likelihood of such phenomenon will help us prevent the occurrence of blast-induced flyrock.

Acknowledgements The authors of this study would like to thank Dr. Armaghani and Dr. Yari for providing the dataset. The first author (R. Zhang) also would like to thank the financial support from Queensland University of Technology and the China Scholarship Council, on this study.

Funding Open Access funding enabled and organized by CAUL and its Member Institutions. No funding was received for conducting this study.

Data availability The orginal data is avaliable upon request to the corresponding author.

Declarations

Conflict of interest The authors have no relevant financial or non-financial interests to disclose.

References

Amini H, Gholami R, Monjezi M, Torabi SR, Zadhesh J (2012) Evaluation of flyrock phenomenon due to blasting operation by support vector machine. Neural Comput Appl 21:2077–2085

Armaghani DJ, Mahdiyar A, Hasanipanah M, Faradonbeh RS, Khandelwal M, Amnieh HB (2016) Risk assessment and prediction of flyrock distance by combined multiple regression analysis and Monte Carlo simulation of quarry blasting. Rock Mech Rock Eng 49:3631–3641

Bajpayee T, Verakis H, Lobb T (1999) An analysis and prevention of flyrock accidents in surface blasting operations. ISEE 2004:401–410

Bhandari S (1997) Engineering rock blasting operations. AA Balkema, Rotterdam

Bisoi R, Dash P, Das PP (2020) Short-term electricity price forecasting and classification in smart grids using optimized multikernel extreme learning machine. Neural Comput Appl 32:1457–1480

Caponnetto A, De Vito E (2007) Optimal rates for the regularized least-squares algorithm. Found Comput Math 7:331–368

Chen L, Asteris PG, Tsoukalas MZ, Armaghani DJ, Ulrikh DV, Yari M (2022) Forecast of airblast vibrations induced by blasting using support vector regression optimized by the grasshopper optimization (SVR-GO) technique. Appl Sci 12(19):9805

Cortes C, Vapnik V (1995) Support-vector networks. Mach Learn 20:273–297

Ding X, Hasanipanah M, Ulrikh DV (2024) Hybrid metaheuristic optimization algorithms with least-squares support vector machine and boosted regression tree models for prediction of air-blast due to mine blasting. Nat Resour Res 2024:1–15

Ghasemi E, Sari M, Ataei M (2012) Development of an empirical model for predicting the effects of controllable blasting parameters on flyrock distance in surface mines. Int J Rock Mech Min Sci 52:163–170

Ghasemi E, Amini H, Ataei M, Khalokakaei R (2014) Application of artificial intelligence techniques for predicting the flyrock distance caused by blasting operation. Arab J Geosci 7:193–202

Guo H, Zhou J, Koopialipoor M, Jahed Armaghani D, Tahir M (2021) Deep neural network and whale optimization algorithm to assess flyrock induced by blasting. Eng Comput 37:173–186

Han H, Jahed Armaghani D, Tarinejad R, Zhou J, Tahir M (2020) Random forest and Bayesian network techniques for probabilistic prediction of flyrock induced by blasting in quarry sites. Nat Resour Res 29:655–667

Harandizadeh H, Armaghani DJ (2021) Prediction of air-overpressure induced by blasting using an ANFIS-PNN model optimized by GA. Appl Soft Comput 99:106904

Hasanipanah M, Faradonbeh RS, Armaghani DJ, Amnieh HB, Khandelwal M (2017) Development of a precise model for prediction of blast-induced flyrock using regression tree technique. Environ Earth Sci 76:1–10

Hasanipanah M, Keshtegar B, Thai D-K, Troung N-T (2022) An ANN-adaptive dynamical harmony search algorithm to approximate the flyrock resulting from blasting. Eng Comput 2022:1–13

Huang G-B, Zhou H, Ding X, Zhang R (2011) Extreme learning machine for regression and multiclass classification. IEEE Trans Syst Man Cybern Part B (Cybern) 42(2):513–529

Hustrulid WA (1999) Blasting principles for open pit mining: general design concepts. Balkema, London

Jahed Armaghani D, Tonnizam Mohamad E, Hajihassani M, Alavi Nezhad Khalil Abad S, Marto A, Moghaddam M (2016) Evaluation and prediction of flyrock resulting from blasting operations using empirical and computational methods. Eng Comput 32:109–121

Jamei M, Hasanipanah M, Karbasi M, Ahmadianfar I, Taherifar S (2021) Prediction of flyrock induced by mine blasting using a novel kernel-based extreme learning machine. J Rock Mech Geotech Eng 13(6):1438–1451

Kloft M, Brefeld U, Sonnenburg S, Zien A (2011) Lp-norm multiple kernel learning. J Mach Learn Res 12:953–997

Koopialipoor M, Fallah A, Armaghani DJ, Azizi A, Mohamad ET (2019) Three hybrid intelligent models in estimating flyrock distance resulting from blasting. Eng Comput 35:243–256

Lin S-B, Guo X, Zhou D-X (2017) Distributed learning with regularized least squares. J Mach Learn Res 18(1):3202–3232

Liu X, Wang L, Huang G-B, Zhang J, Yin J (2015) Multiple kernel extreme learning machine. Neurocomputing 149:253–264

Lundborg N, Persson A, Ladegaard-Pedersen A, Holmberg R (1975) Keeping the lid on flyrock in open-pit blasting. Eng Min J 176:95–100

Marto A, Hajihassani M, Jahed Armaghani D, Tonnizam Mohamad E, Makhtar AM (2014) A novel approach for blast-induced flyrock prediction based on imperialist competitive algorithm and artificial neural network. Sci World J 2014:1

Mohamad ET, Armaghani DJ, Hajihassani M, Faizi K, Marto A (2013) A simulation approach to predict blasting-induced flyrock and size of thrown rocks. Electron J Geotech Eng 18(B):365–374

Monjezi M, Amiri H, Farrokhi A, Goshtasbi K (2010) Prediction of rock fragmentation due to blasting in Sarcheshmeh copper mine using artificial neural networks. Geotech Geol Eng 28:423–430

Monjezi M, Khoshalan HA, Varjani AY (2011) Optimization of open pit blast parameters using genetic algorithm. Int J Rock Mech Min Sci 48(5):864–869

Monjezi M, Amini Khoshalan H, Yazdian VA (2012) Prediction of flyrock and backbreak in open pit blasting operation: a neuro-genetic approach. Arab J Geosci 5(3):441–448

Monjezi M, Mehrdanesh A, Malek A, Khandelwal M (2013) Evaluation of effect of blast design parameters on flyrock using artificial neural networks. Neural Comput Appl 23:349–356

Monjezi M, Dehghan JAH, Samimi NF (2007) Application of TOPSIS method in controlling fly rock in blasting operations. In: proceedings of seventh international science conference SGEM. Sofia, Bulgaria, pp 41–41

Morris MD (1991) Factorial sampling plans for preliminary computational experiments. Technometrics 33(2):161–174

Murlidhar BR, Nguyen H, Rostami J et al (2021) Prediction of flyrock distance induced by mine blasting using a novel Harris Hawks optimization-based multi-layer perceptron neural network. J Rock Mech Geotech Eng 13(6):1413–1427

Newcombe RG (1998) Two-sided confidence intervals for the single proportion: comparison of seven methods. Stat Med 17(8):857–872

Nguyen H, Bui X-N, Choi Y, Lee CW, Armaghani DJ (2021) A novel combination of whale optimization algorithm and support vector machine with different kernel functions for prediction of blasting-induced fly-rock in quarry mines. Nat Resour Res 30:191–207

Olofsson SO (1990) Applied explosives technology for construction and mining. Applex, London

Rad HN, Hasanipanah M, Rezaei M, Eghlim AL (2018) Developing a least squares support vector machine for estimating the blast-induced flyrock. Eng Comput 34:709–717

Raina A, Chakraborty A, Ramulu M, Sahu P, Haldar A, Choudhury P (2004) Flyrock prediction and control in opencast mines: a critical appraisal. Min Eng J 6(5):10–20

Rezaei M, Monjezi M, Varjani AY (2011) Development of a fuzzy model to predict flyrock in surface mining. Saf Sci 49(2):298–305

Saghatforoush A, Monjezi M, Shirani Faradonbeh R, Jahed AD (2016) Combination of neural network and ant colony optimization algorithms for prediction and optimization of flyrock and back-break induced by blasting. Eng Comput 32:255–266

Saunders C, Gammerman A, Vovk V (1998) Ridge regression learning algorithm in dual variables. In: Proceedings of the fifteenth international conference on machine learning. Morgan Kaufmann. pp 515–521

Trivedi R, Singh T, Raina A (2014) Prediction of blast-induced flyrock in Indian limestone mines using neural networks. J Rock Mech Geotech Eng 6(5):447–454

Trivedi R, Singh T, Gupta N (2015) Prediction of blast-induced flyrock in opencast mines using ANN and ANFIS. Geotech Geol Eng 33:875–891

Yari M, Monjezi M, Bagherpour R (2013) Selecting the most suitable blasting pattern using AHP-TOPSIS method: Sungun copper mine. J Min Sci 49:967–975

Yari M, Monjezi M, Bagherpour R, Jamali S (2014) Developing a mathematical assessment model for blasting patterns management: Sungun copper mine. J Cent South Univ 21:4344–4351

Yari M, Bagherpour R, Jamali S (2017) Development of an evaluation system for blasting patterns to provide efficient production. J Intell Manuf 28:975–984

Yari M, Armaghani DJ, Maraveas C, Ejlali AN, Mohamad ET, Asteris PG (2023) Several tree-based solutions for predicting flyrock distance due to mine blasting. Appl Sci 13(3):1345

Zhang R, Li Y, Gui Y, Zhou J (2022) Prediction of blasting induced air-overpressure using a radial basis function network with an additional hidden layer. Appl Soft Comput 127:109343

Zhang R, Li Y, Gui Y (2023) Prediction of rock blasting induced air overpressure using a self-adaptive weighted kernel ridge regression. Appl Soft Comput 148:110851

Zhang R, Li Y, Gui Y, Armaghani DJ, Yari M (2024a) A stacked deep multi-kernel learning framework for blast induced flyrock prediction. Int J Rock Mech Min Sci 177:105741

Zhang R, Li Y, Gui Y, Armaghani DJ, Yari M (2024b) A stacked multiple kernel support vector machine for blast induced flyrock prediction. Geohazard Mech 2024:1