Shahrood University of Technology

Iranian Society of Mining Engineering (IRSME)

Journal of Mining and Environment (JME)

journal homepage: www.jme.shahroodut.ac.ir

THOMSON REUTERS

A State of the art Catboost-Based T-Distributed Stochastic Neighbor Embedding Technique to Predict Back-break at Dewan Cement Limestone Quarry

Muhammad Kamran*

Department of Mining Engineering, Institute Technology of Bandung, Indonesia

Article Info

Received 17 September 2021

Received in Revised form 27 September 2021

Accepted 29 September 2021

Published online 29 September 2021

DOI:10.22044/jme.2021.11222.2104

Keywords

Blasting

Back-Break

Catboost

Rock engineering

Mining industry

Abstract

The blasting operation is an important rock fragmentation technique employed in several foundation engineering disciplines such as mining, civil, tunneling, and road planning. Back-break (BB) is one of the adverse effects caused by the blasting operations that produces several effects including vulnerability of mining machinery, bench slope design, and risks to the next blast-patterns due to the eruption of gases from several discontinuities in jointed rock masses. Several techniques have been executed by the researchers in order to predict BB in the blasting operations. However, this is the first work to implement a-state-of-the-art Catboost-based t-distributed stochastic neighbor embedding (t-SNE) approach to predict BB. A total of 62 datasets having 12 influential BB-generating features are collected from genuine blasting patterns. A novel dimensionality depletion technique t-SNE that operates the Kullback-Leibler divergence interpretation is employed to tailor the pioneer exaggeration of the blasting dataset. Then the t-SNE dataset obtained is split into a 70:30 ratio of the training and testing datasets. Finally, the Catboost method is implemented on a low-dimensionality blasting database. The performance evaluation criterion confirms that the BB predictive model is more stable with a goodness of fit = 99.04 in the training dataset, 97.26 in the testing datasets, and could anticipate a more accurate prediction. Moreover, the model presented in this work performs superior to the existing publicly available execution of BB. In summary, this model can be practiced in order to predict BB in several rock engineering practices and mining industry scenarios.

1. Introduction

Blasting is one of the significant rock fragmentation techniques employed in several engineering disciplines including mining, civil, tunneling, and road planning. In order to reduce overall cost of blasting operation, enhancing the drilling process, increasing the price of loading and shipment of bulk materials, and improving the effectiveness after mineral's extraction, the blasting engineers have to deal with a desired blasting operation [1]. In a blasting operation, an abundant portion of explosive energy is dissipated to several environment effects including back-break (BB), ground vibrations, fly-rock, and air-overpressure, which can influence the adjacent locality [2-6]. The fragmented rock beyond the boundary of rear row of drill holes in a blasting design is known as BB. BB causes the vulnerability to the mining machinery and bench slope blasting [7], and may interrupt the mining operation.

The control blasting techniques including line drilling, cushion blasting, and pre-shearing have been introduced to reduce BB [8]. However, these methods are typically time-consuming and expensive [9]. Various researchers have proposed several blast design features causing BB. An adaptable and robust framework is acquired to obtain an ideal blast pattern [10]. The traditional statistical methods may mislead to obtain BB

Corresponding author: muhdkamran62@gmail.com (M. Kamran).

generating cause due to its various non-linear attributes [11].

BB can produce some additional risks to the next blast-patterns considering the breakout of gases from various discontinuities within rock mass [7]. In order to analyse the environmental effects of blasting, various artificial intelligence techniques have been developed by the researchers [12-21].

Khandelwal and Singh [22] have used an artificial neural network in order to establish a new framework on ground vibration prediction resulting from blasting operations. A fly-rock distance has been predicted by Rezaei et al. [23]. The model results in depict that the developed fuzzy model predicts the fly-rock with a high accuracy as compared to the traditional statistical models. Various machine learning algorithms including artificial neural network, fuzzy model, and regression frameworks have been utilized to predict BB at the Sangan iron mine Iran dataset [24]. Their study reveals that as compared to the other frameworks, the ANFIS model has a high efficiency to forecast BB. A novel mechanism was established based on the hybrid particle swarm optimization (PSO) and artificial neural network in order to predict the air blast-overpressure [25]. Their results depict that the presented technique predict the airblast-overpressure with a high accuracy. Table 1 presents the recent up-to-model artificial intelligence models to predict BB [26-35].

Table 1. Advanced artificial intelligence framework forecasting BB.

AI approaches	Input attributes	Output attribute	No of Patterns	R^2	RMSE	References
GP	B, S, ST, PF, SR	BB	175	0.98	0.32	16
ANN	B, S, ST, L, PF	BB	34	0.77	0.53	17
ANN	B, S, ST, N, PF, DPM, SD, RF	BB	–	0.86	0.49	18
ANN	B, S, L, ST, PF, SD	BB	103	0.87	0.22	19
SVM	B, S, L, SD, ST, PF	BB	193	0.92	0.34	20
ANN, ANFIS	SR, ST, PF, RD, N, CLR, S/B	BB	42	ANN = 0.92 ANFIS = 0.96	ANN = 0.88 ANFIS = 0.6	21
ANN	UCS, SD, WC, B, S, ST, D, BH, PF, C	BB	97	0.9	–	22
GA-ANN	D, L, B, S, ST, PF, SD, C, RMR	BB	195	0.96	–	23
FIS	B, S, ST, SD, PF, HD, C, RD	BB	–	0.95	–	24
ANN	B, CLR, PF, S/B, ST/B, N	BB	300	–	0.64	25

where S = spacing, B = burden, PF = powder factor, ST = stemming, N = number of rows, C = charge per delay, L = hole depths, SR = stiffness ratio, SD specific drilling, GP = genetic programming, ANN = artificial neural network, RD = rock density, CLR = last row charge per total charge ratio, Q = linear charge concentration, WC = water content, DPM = delay per meter, RF = rock factor, BH = bench height.

Based on the above-mentioned literature, the implementation of the state-of-the-art Catboost-based T-distributed stochastic neighbor embedding (t-SNE) has not been employed to predict BB. In this work, an attempt is introduced to predict BB associated with a blasting operation by employing the blasting database at Dewan Cement Limited, Pakistan. This work introduces a state-of-the-art unsupervised data depletion technique t-SNE to tailor the exaggeration of the blasting dataset, and later, a supervised machine learning algorithm i.e. Catboost to predict BB in the blasting operations. To the best of the author information, this is the first work to execute the state-of-the-art t-SNE+Catboost method to predict BB. A flowchart of this work is elucidated in Figure 1.

2. Data acquisition

The dataset acquired in this work is from the published article of Dewan cement limestone quarry, Pakistan [11]. The Dewan cement limestone quarry is located in the Khyber Pakhtunkhwa Province in Pakistan at a distance of 4.5 km from the main crushing plant [11]. The main rock constituent at Dewan Cement Limited is sedimentary rock, i.e. at the upper plate is limestone and the bottom plate is siltstone-claystone to claystone. The typical perspective view of the Dewan cement limestone quarry is shown in Figure 2. Blasting is the main operation employed to fragment the rocks in order to provide the raw material for manufacturing cement. The overall dataset consists of 62 blast patterns having 12 influential attributes including burden, spacing, hole depth (m), blasthole inclination, high explosive %, ANFO%, Stemming (m), powder factor, delay per row, hole-to-hole delay, No. of rows, No. of free faces, and the corresponding BB. The statistical description of the blasting patterns is shown in Table 2. The Python software with the Seaborn module is utilized to present the pairwise correlation of original blasting database. Figure 3 shows the pairwise correlation of various influential attributes to BB. From this figure, it is clear that burden, hole depth (m), blast hole inclination, ANFO%, stemming (m), powder factor, hole-to-hole delay, and number of rows are positively correlated to BB, whereas spacing, high explosive %, delay per row, and number of free faces are negatively correlated to BB. Moreover, it can also be concluded from Figure 3 that a bunch of attributes are relatively low correlated to BB, hence all the attributes are considered in order to improve the accuracy of the model.

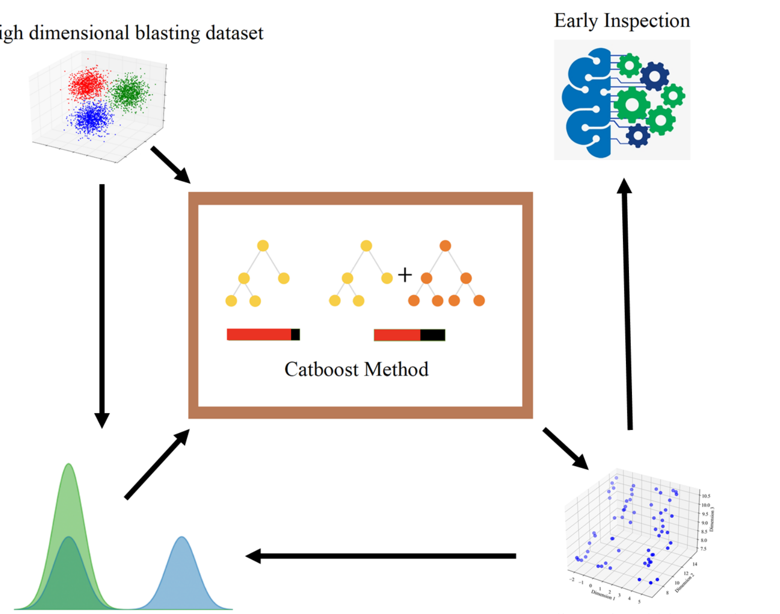

Figure 1. A flowchart of the work.

Figure 2. A typical perspective view of Dewan cement limestone quarry.

Table 2. Statistical description of blasting patterns.

	Burden	Spacing	Hole depth (m)	Blast hole inclination	High explosive %	ANFO%	Stemming (m)	Powder factor	Delay per row	Hole-to-hole delay	No. of rows	No. of free faces	Back-break BB (m)
Mean	3.1	4.09	11.76	78.3	16.04	83.95	2.49	1.03	33.06	33.87	1.45	1.79	2.37
Std	0.2	0.26	3	4.43	4.14	4.14	0.3	0.35	32.88	15.07	0.61	0.65	2.29
Min	2.5	3.5	6.09	70	8	70.59	2	0.5	0	25	1	1	0
Max	4	4.5	16.16	85	29.41	92	3	1.4	100	75	3	3	8

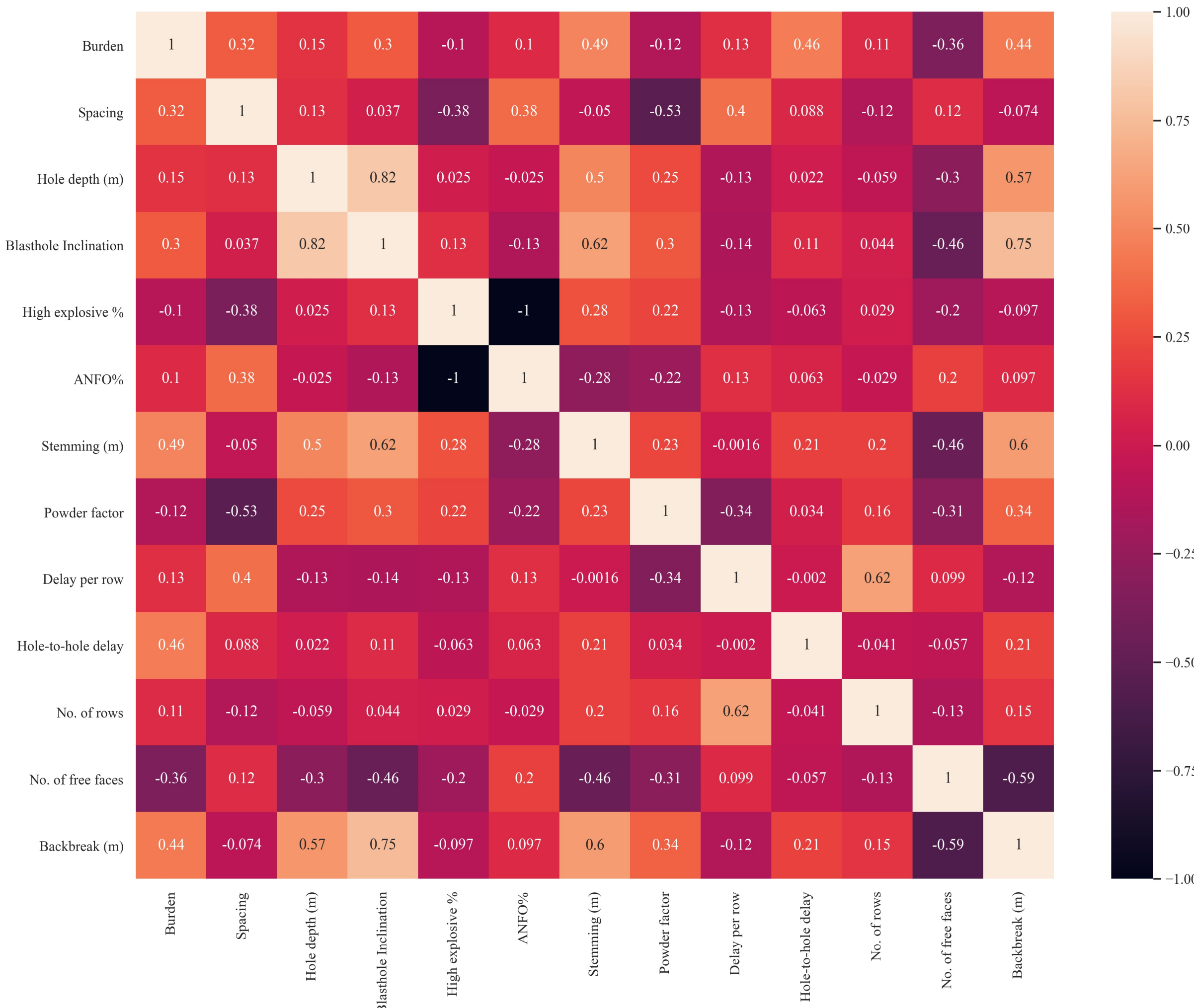

Figure 3. Pairwise correlation plot for robust attributes selection in blasting dataset.

3. Methodology

3.1. Distributed stochastic neighbor embedding (t-SNE) technique overview

T-distributed Stochastic Neighbor Embedding (t-SNE) is a state-of-the-art deep learning algorithm employed for non-linear dimension depletion in order to process the data visualization [36-37]. T-SNE is an updated modification of stochastic neighbor embedding (SNE) introduced by Van der Maaten and Hinton [38-39]. This technique achieves optimum consequences for data visualization, which permits an extrapolation of the points in a low dimensional region that excels it easy to utilize for a large set of data [40]. The parameter t-SNE includes learning rate that affects the speed at which the model tends or moves toward one point or one another in back-propagation [41]. The T-SNE technique has been widely employed in the rock engineering studies in order to visualize high-dimensional datasets [42-44].

3.2. Catboost description

Catboost is one of the significant gradient boosted machine learning frameworks introduced by Dorogush et al. in the recent time [45]. This algorithm can handle both the classification and regression complication, and has been published as an open access full-featured gradient boosting archive [45-46]. The execution time and memory manipulation of Catboost are far fewer than the other machine learning algorithms [47]. While handling the categorical attributes, Catboost could be a better replacement with superlative results [48]. In contrast, with the point to point training of continuous attributes by the machine learning algorithms, Catboost locates the mixed attributes with a high precision in rock mechanics prediction models [49].

3.3. Blasting dataset visualization employing t-SNE

3.3.1. Stochastic neighbor embedding (SNE)

The t-SNE implemented in this work is an advanced dimensionality depletion unsupervised machine learning technique contingent on the stochastic neighbor embedding. SNE converts the euclidean distance in a multi-scale data point to conditional uncertainty in order to reveal the resemblance among the data points. In a dataset, the conditional uncertainty $U_{n|m}$ is employed to constitute the resemblance of point a_m, which is given in Equation 1.

$$U_{n|m} = \begin{cases} \dfrac{\exp(-||a_m - a_n||^2/2{\sigma_i}^2)}{\sum_{k\neq 1} \exp - ||a_m - a_n||^2/(2\sigma_i)^2} \\ 0 \;\; m = n \end{cases} \quad (1)$$

where $m \neq n$

Where σ_i depicts the Gaussian distribution variance having a_m as the center position, which is established by a binary search by employing the mechanism of perplexity. The perplexity is shown in Equation 2.

$$Perp(U_m) = 2^{H(U_i)} \quad (2)$$

where, $E(U_m)$ is the entropy of U_i, as given in Equation 3.

$$E(U_m) = -\sum U_{n|m} log_2 U_{n|m} \quad (3)$$

Suppose that b_m and b_n are locality in the low dimension that are designated to a_m and a_n in the high dimension. In our case, the Gaussian distribution is particularizing as $\frac{1}{\sqrt{2}}$. Succeeding the resemblance of $U_{n|m}$ of a_n to a_m is given in Equation 4.

$$U_{n|m} = \begin{cases} \dfrac{\exp(-||a_m - a_n||^2)}{\sum_{k\neq 1} \exp(-||a_m - a_n||^2)} \\ 0 \;\; m = n \end{cases} \quad (4)$$

$m \neq n$

If the dimensionality depletion outcome is satisfactory, then the resemblance in a high dimensionality space is assumed to be identical to that in low dimensionality i.e. $U_{n|m} = V_{n|m}$. When the conditional uncertainty between a_m and all the other points are examined, the conditional uncertainty distribution U_n can be established. Correspondingly, the identical uncertainty distribution V_n is established as the U_n low dimensionality space. In order to measure the resemblance between two points, the Kullback-Leibler divergence is employed. Hence, a cost function F is established, as shown in Equation 5.

$$F = \sum_i KL(U_m||V_m) = \sum_m \sum_n U_{n|m} log \frac{U_{n|m}}{V_{n|m}} \quad (5)$$

The gradient formula is defined by Equation 6.

$$\frac{\delta C}{\delta y_i} = 2 \sum_n \left(U_{n|m} - V_{n|m} + U_{m|n} - V_{m|n}\right)(a_m - a_n) \quad (6)$$

3.3.2. t-SNE based on SNE

The t-SNE technique is an updated mechanism calculated on the SNE technique, which utilizes the joint distribution in place of conditional one, as shown in Equation 7.

$$U_{mn} = \frac{U_{m|n} + U_{n|m}}{2} \quad (7)$$

The t-distribution is the important dissimilarity between the SNE and t-SNE techniques in preference to the Gaussian method. Rather, Gaussian is employed in a low dimensionality region to transform a data into an uncertainty distribution, whereas Gaussian is still employed in a high dimensionality region, and when the freedom degree is 1, then V_{mn} is defined by Equation 8.

$$V_{mn} = \frac{1 + (||a_m - a_n||^2)^{-1}}{\sum_{k\neq 1}(1 + ||a_n - a_o||^2)^{-1}} \quad (8)$$

Finally, the optimized gradient is given by Equation 9.

$$\frac{\delta C}{\delta y_i} = 4 \sum_n \left(U_{n|m} - V_{n|m} + U_{m|n} - V_{m|n}\right)(a_m - a_n) \quad (9)$$

In order to make it more precise, the comprehensive mechanism of t-SNE is given as:

Stage 1: Get data U= $U_1, U_2, U_3, \ldots, U_n$ in a high dimension region, and assign the dimensionality depletion consequences as $B^{(T)} = V_1, V_2, V_3, \ldots, V_n$.

Stage 2: Compute perplexity, and assign iteration times *T*, momentum of $\alpha(t)$, and learning rate η.

Stage 3: Calculate $U_{m|n}$, as given in Equation 1.

Stage 4: Estimate U_{mn}, as depicted in Equation 7.

Stage 5: Arbitrary choose Y with N.

Stage 6: Compute V_{mn} as stated in Equation 8, and estimate the gradient as stated in Equation 9.

Stage 7: Finally repeat the stage 6 procedure so that the iteration number is remarkable than T.

3.4. BB prediction employing Catboost method

In the Catboost method, the decision tree is utilized as a weak learner. The Categorical and text attributes of a dataset can be handled with Catboost without the executor required to tackle them individually. It also anticipates grid search and randomized search, which assist in finding out a catalogue of attributes value to search the best composition of attributes that permits the optimum outcome. The default attributes settings in the Catboost library usually gives a good fitness model. Besides, Catboost also anticipates assistance for executing the training stage on GPU and tenfold hyper-parameter optimization with straight forward alignment to improve the diverse practicable rock and soil mechanics circumstances.

The Catboost method reduces the prediction relocation that occurs during the training stage. The prediction relocation is the elimination of $\mathbf{F(a)|}\ (\boldsymbol{a_i})$ with $\boldsymbol{a_i}$ being *a* training instance, in association with $\mathbf{F(b)|}\ (\boldsymbol{b_i})$ being a test instance *b*. The gradient boosting utilizes the same instance for the estimation of gradient, and the model minimizes the gradient. The idea of Catboost is to build a base model for the discrete *D* boosting iterations. The *ith* model of the *mth* iterations is trained on the *ith* instance of the permutation, and is appropriate to estimate the gradient of *j+1* instance for *p+1* repetition. Moreover, this mechanism utilizes *r* reciprocated inconsistent presentation. Finally, an identical model is built per iteration that contains all the framework permutations. Following the same splitting criteria, the tree is outstretched by fattening the leaf nodes level-wise.

The technique implemented in Catboost is to estimate the contemporary attribute resemblance to the one that mimics for building the network. Hence, for a specified inconsistent permutation of the sample, the data sample $< i$ is employed to the attribute value for each discrete sample *i*. Finally, following several permutations in the blasting dataset, the acquired attributes are averaged. Figure 4 depicts the Catboost method implemented in this work.

It is important to mention that the Catboost method training capability is governed by its model hypermeters including the iteration number, maximum depth and learning rate, etc. The designation of the optimal hyper-parameters for a model is a laborious, complicated, and challenging task; however, it is based on the executor competences and expertise.

4. Results and discussion

The Jupyter notebook, a programing environment, is an open-source implementation particularly employed in engineering for establishing and sharing the scholastic ideas, incorporating various categories of resources including images, texts, and codes in different programing languages in an isolated catalogue, approachable through a web portal. The Jupyter notebook is also appropriate in order to provide access to an online data analysis and elucidating how to manipulate the data [50]. The Python software is utilized in the Jupyter notebook to accomplish t-SNE with the Scikit-learn module [51]. In the first step, the blasting dataset is visualized from a high-dimensional space to a low-dimensional space. The original blasting dataset is categorized into three clusters. In this work, the drill hole attributes including the burden, spacing, hole depth (m), blast hole inclination, stemming (m), powder factor, No. of rows, and No. of free faces are considered in the first cluster (Dimension 1). The delay time attributes including the delay per row and the hole-to-hole delay are categorized in the second cluster (Dimension 2). The explosive types, i.e. high explosive % and ANFO% are grouped in the third cluster (Dimension 3). The *learning rate = 100* is selected in the Python programming language with the Matplotlib module in order to visualize the original dataset (all the other parameters are kept as a default). Following the blasting dataset depletion mechanism, the attributes demonstrating space are kept with an unsophisticated visualization in such a way that the original blasting features may retain the originality to the maximum extent. Table 3 elucidates the blasting dataset after dimensionality depletion. Finally, following the t-SNE approach, the actual blasting dataset (62 × 12 matrix) is transformed to a (62 × 3) matrix, as shown in Table 3. Figure 5 shows a 3D space of blasting dataset after the t-SNE data depletion approach.

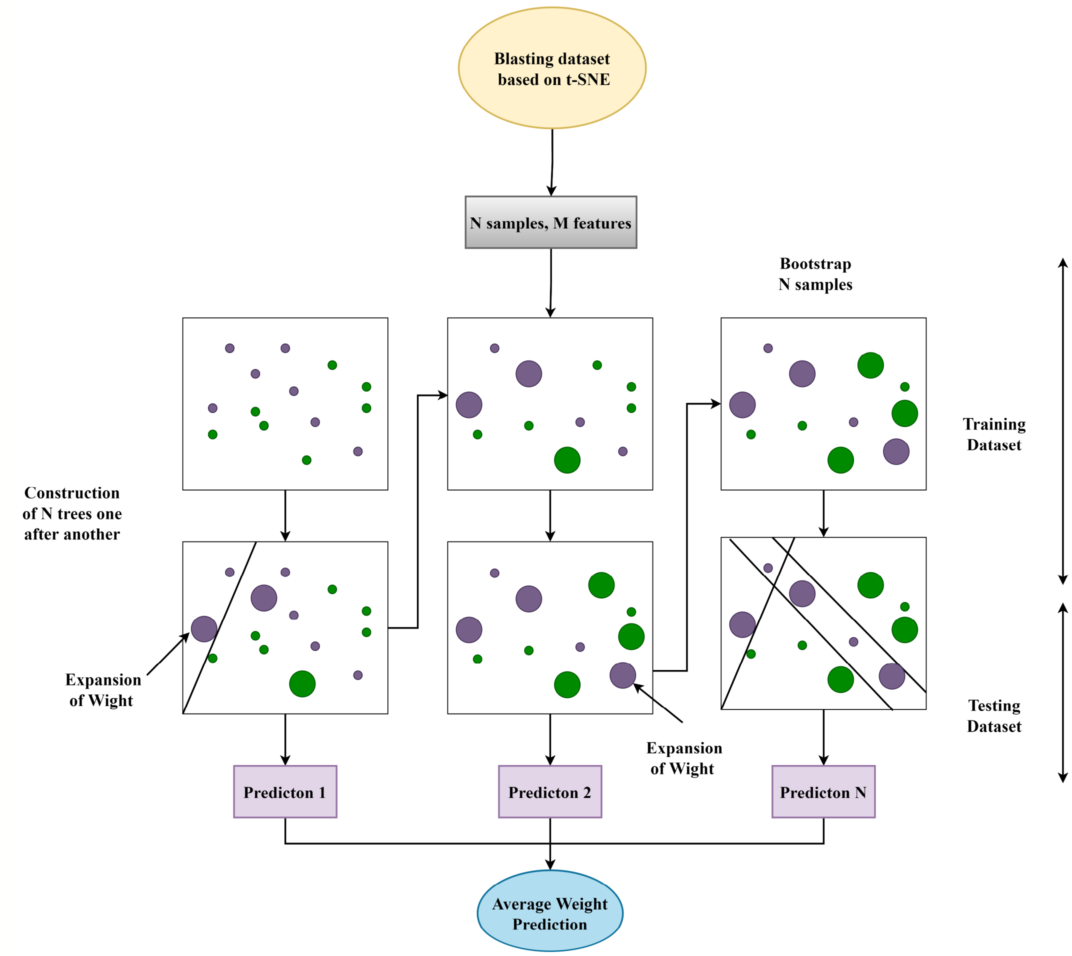

Figure 4. Catboost method interpretation implemented in this work.

Table 3. Data after attribute reduction with t-SNE approach.

Blasting patten No	Dimension 1	Dimension 2	Dimension 3
1	-2.0602772	6.6390648	7.7710314
2	-2.0596974	6.4246774	7.9374428
3	-2.0411243	7.9593358	7.588429
4	0.1286853	7.9593434	7.588428
5	-2.1744084	7.3905997	7.5189633
…..	…..	…..	…..
58	2.8748353	12.374563	9.4801922
59	2.1663244	14.227405	10.563422
60	3.6073992	11.243856	9.2344484
61	2.8757045	13.689963	10.576619
62	3.764091	12.977768	9.9173899

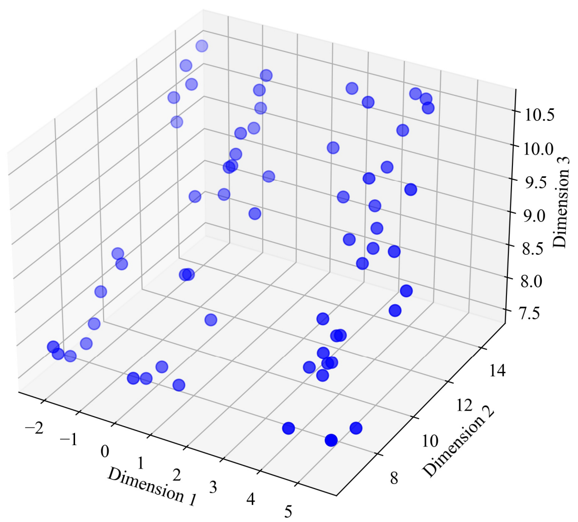

Figure 5. 3D space of blasting dataset after t-SNE data depletion approach.

In the second step, the low-dimensional space and the actual BB are combined in order to make a dataset for the Catboost method. In order to build the machine learning model, the overall dataset is utilized to build the model [52]. Following the same mechanism, Catboost is developed by an entire t-SNE dataset. Later on, the data obtained from the t-SNE model is split into the training and testing datasets. Hence, the real time monitoring of the next inspector cycle is estimated by Catboost. Finally, the prediction accuracy is evaluated both in the training and testing datasets.

The t-SNE data obtained is distributed into 70% (43 patterns) for training the Catboost, and the remaining 30% (19 patterns) is employed for the testing purposes. The hyperparameter, i.e. *verbose=80, n_estimators=500* is selected in the Python programming language using the Jupyter notebook to execute Catboost. (All the other parameters are kept as a default.)

Several indices are used by the researchers in order to measure the robustness of regression based the machine learning algorithms [53-55]. This work computed the prediction execution of Catboost from five different aspects including MSE, RMSE, MAE, VAF, and R^2.

MSE is an index to evaluate the error between the spatial feature BB and the actual BB, which is defined by the Equation 10.

$$MSE = \frac{1}{T}\sum_{i=1}^{n}(BB_{actual} - BB_{mean})^2 \quad (10)$$

RMSE is another regression evaluation metric shown by Equation 11.

$$RMSE = \sqrt{\frac{1}{T}\sum_{i=1}^{n}(BB_{actual} - BB_{mean})^2} \quad (11)$$

MAE is elucidated by Equation 12.

$$MAE = \sum_{i=1}^{n}|\ BB_{mean} - BB_{actual}| \quad (12)$$

Value account For (VAF) is another significant index used to estimate how powerful a regression model is in a rock mechanics model [56], and is given by Equation 13.

$$VAF = \left[1 - \frac{var(BB_{actual} - BB_{SFD})}{var(BB_{actual})}\right] \times 100 \quad (13)$$

In order to measure the fitness of the model, coefficient of correlation R^2 is implemented in this article, which is given by Equation 14.

$$R^2 = \left[1 - \frac{\sum_{i=1}^{n}(BB_{actual} - BB_{SFD})^2}{\sum_{i=1}^{n}(BB_{actual} - BB_{mean})^2} \times 100\right] \quad (14)$$

where BB_{SFD} and BB_{mean} depict the spatial feature BB data predicted by Catboost, and the mean BB data respectively, T is the total number of blasting datasets, and BB_{actual} is the actual BB value.

Table 4 depicts the statistical results of the Catboost model. The results tabulated in Table 2 affirm that the Catboost method predicts the spatial feature BB with lower RMSE, MSE, and MAE, and with a high VAF and goodness of fit in both the training and testing datasets.

Table 4. Regression evaluation matrix of Catboost method.

	MSE	RMSE	MAE	VAF	Goodness of fit
Training dataset	0.068	0.260	0.207	99.078	99.049
Testing dataset	0.080	0.282	0.224	97.585	97.260

Figures 6 and 7 depict the scattered plots of the actual versus spatial feature BB predicted by the Catboost method at the training and testing stages, respectively. In addition, in order to realize the presentation of the spatial feature BB combined with the actual BB, Figures 8 and 9 present the performance of the Catboost prediction of the spatial feature BB against the actual values at the (a) training data (b) testing data.

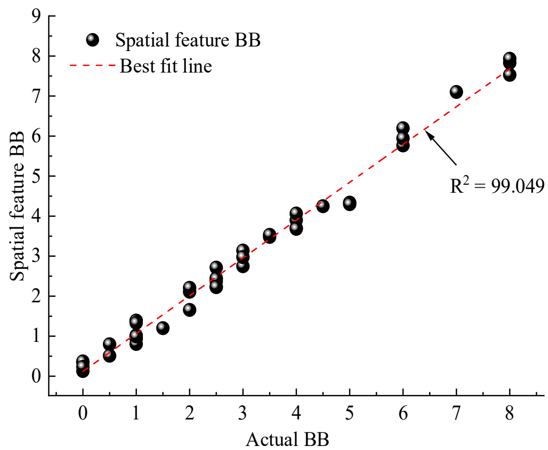

Figure 6. Catboost prediction against actual BB at training dataset.

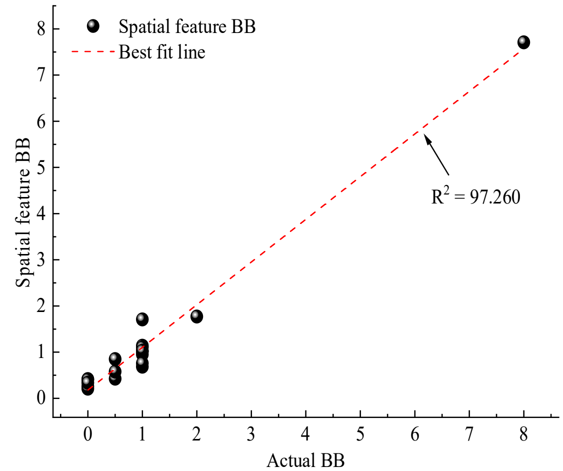

Figure 7. Catboost prediction against actual BB at testing dataset.

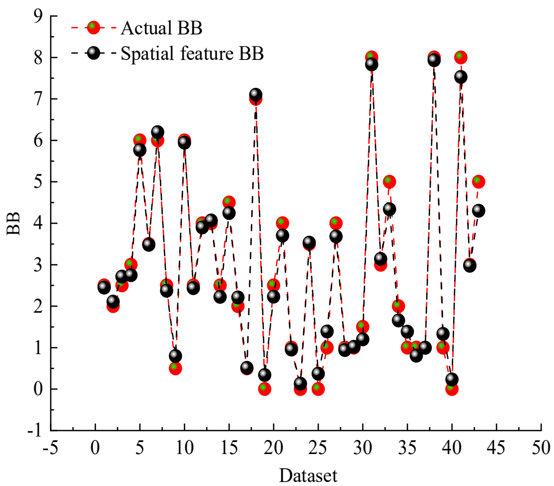

Figure 8. Performance of Catboost prediction versus actual values at training dataset.

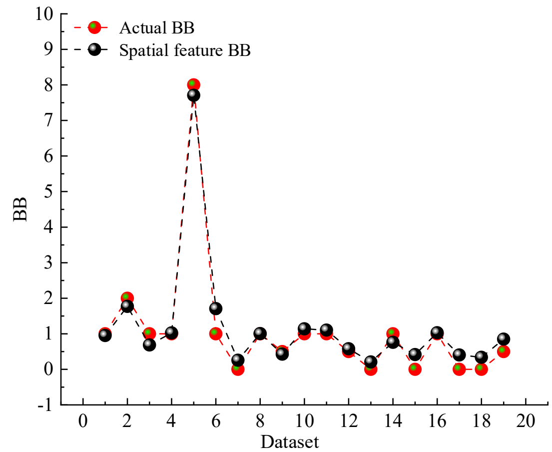

Figure 9. Performance of Catboost prediction versus actual values at testing dataset.

The Catboost-based t-SNE model estimates a high fitting degree. The t-SNE model adopted in this work utilizes the non-linear dimensionality depletion technique to reserve the local and overall characteristics of the non-linear blasting data, and then utilizes Catboost to map the non-linear relationship between the blasting data and BB. Hence, the model exhibits a high prediction accuracy in order to predict BB as compared to the existing publicly available literature [11, 26-35]. Hence, the model suggested in this article can be implemented in several rock engineering practices and mining industry scenarios.

5. Conclusions

In this work, the author proposed a two-step state of the art technique based on t-SNE+Catboost for the prediction of BB. The t-SNE mechanism adopted in this article retains the local and overall attributes of a high-dimensional dataset, and employs the Catboost method to mimic the non-linear blasting monitoring database. The model manifests a high performance in predicting BB in reference to the lists of the published articles.

It can be obtained that the BB prediction model based on t-SNE+Catboost has a MSE = 0.068, RMSE = 0.260, MAE = 0.207, VAF = 99.070, goodness of fit of 99.049 in the training dataset, and MSE = 0.080, RMSE = 0.282, MAE = 0.224, VAF = 97.585, and goodness of fit of 99.040 in the testing dataset. The blasting engineers should employ the proposed artificial intelligence (AI) model to predict back-break because the artificial intelligence model learns and improves the model performance of the blasting patterns by using data and experience.

The author future work will focus on predicting the blasting side-effect by metaheuristic techniques in order to determine the optimized blasting pattern with a minimal back-break and enhancing the accuracy of the model prediction and the model performance in heterogeneous and big datasets. Moreover, the author plans to investigate BB by using the optimized-based machine learning algorithms, hybrid and ensemble learning. Besides, some other influential attributes will be added to the blasting dataset that will robust the BB risk by employing the predicted outcomes of the models.

6. Acknowledgments

The author wishes to express his deep and sincere gratitude to Dr. Barkat Ullah (School of Resources and Safety Engineering, Center South University, Changsha, China) for reading the initial version of the article and his fruitful discussion. Many thanks also goes to Engr. Qadir Bakhsh (Assistant Manager Quarry Department, Dewan Cement Limited) for providing a typical perspective view of the Dewan Cement Limestone Quarry.

References

[1]. Bhandari, S. (1997). Engineering rock blasting operations.

[2]. Raina AK, Murthy VMSR, Soni AK. (2014) Fly-rock in bench blasting: a comprehensive review. Bull Eng Geol Environ. doi:10.1007/s10064-014-0588-6.

[3]. Raina, A.K., Haldar, A., Chakraborty, A.K., Choudhury, P.B., Ramulu, M., and Bandyopadhyay, C. (2004). Human response to blast-induced vibration and air-overpressure: an Indian scenario. Bulletin of Engineering Geology and the Environment. 63 (3): 209-214.

[4]. Marto, A., Hajihassani, M., Jahed Armaghani, D., Tonnizam Mohamad, E., and Makhtar, A.M. (2014). A novel approach for blast-induced fly-rock prediction based on imperialist competitive algorithm and artificial neural network. The Scientific World Journal, 2014.

[5]. Hajihassani, M., Armaghani, D.J., Monjezi, M., Mohamad, E.T., and Marto, A. (2015). Blast-induced air and ground vibration prediction: a particle swarm optimization-based artificial neural network approach. Environmental Earth Sciences. 74 (4): 2799-2817.

[6]. Armaghani, D.J., Mohamad, E.T., Hajihassani, M., Abad, S.A.N.K., Marto, A., and Moghaddam, M.R. (2016). Evaluation and prediction of fly-rock resulting from blasting operations using empirical and computational methods. Engineering with Computers. 32 (1): 109-121.

[7]. Rustan, A. (1998). Rock blasting terms and symbols.

[8]. Wyllie D.C., Mah C., 2004. Rock Slope Engineering, Fourth Edition: Fourth edition. Taylor and Francis. Retrieved from http://books.google.com.pk/books?id=4Gd7Hg2tz-sC.

[9]. Workman L., 1992. Wall Control. Retrieved from http://intrawww.ing.puc.cl/siding/public/ingcursos/cursos_pub/descarga.phtml?id_curso_ic=1781&id_archivo=69274.

[10]. Monjezi M. and Dehghani H., 2008. Evaluation of effect of blasting pattern parameters on back-break using neural networks. International Journal of Rock Mechanics and Mining Sciences. 45 (8):1446-1453. http://doi.org/10.1016/j. ijrmms.2008.02.007.

[11]. Muhammad, K. and Shah, A. (2017). Minimizing back-break at the Dewan Cement limestone quarry using an artificial neural network. Archives of Mining Sciences. 62 (4).

[12]. Kamali, M. and Ataei, M. (2010). Prediction of blast induced ground vibrations in Karoun III power plant and dam: a neural network. Journal of the Southern African Institute of Mining and Metallurgy. 110 (8): 481-490.

[13]. Ghasemi, E., Sari, M., and Ataei, M. (2012). Development of an empirical model for predicting the effects of controllable blasting parameters on fly-rock distance in surface mines. International Journal of Rock Mechanics and Mining Sciences. 52: 163-170.

[14]. Ghasemi, E., Ataei, M., and Hashemolhosseini, H. (2013). Development of a fuzzy model for predicting ground vibration caused by rock blasting in surface mining. Journal of Vibration and Control. 19 (5): 755-770.

[15]. Taji, M., Ataei, M., Goshtasbi, K., and Osanloo, M. (2013). ODM: a new approach for open-pit mine blasting evaluation. Journal of vibration and control. 19 (11): 1738-1752.

[16]. Ataei, M. and Kamali, M. (2013). Prediction of blast-induced vibration by adaptive neuro-fuzzy inference system in Karoun 3 power plant and dam. Journal of Vibration and Control. 19 (12): 1906-1914.

[17]. Ghasemi, E., Amini, H., Ataei, M., and Khalokakaei, R. (2014). Application of artificial intelligence techniques for predicting the fly-rock distance caused by blasting operation. Arabian Journal of Geosciences. 7 (1): 193-202.

[18]. Sari, M., Ghasemi, E., and Ataei, M. (2014). Stochastic modeling approach for the evaluation of back-break due to blasting operations in open pit mines. Rock Mechanics and Rock Engineering. 47 (2): 771-783.

[19]. Ataei, M. and Sereshki, F. (2017). Improved prediction of blast-induced vibrations in limestone mines using Genetic Algorithm. Journal of Mining and Environment. 8 (2): 291-304.

[20]. Hoseini, S.M., Sereshki, F., and Ataei, M. (2018). A quantitative model for evaluation and classification of

blastings in open-pit mines. Journal of Mining and Environment. 9 (1): 127-141.

[21]. Mottahedi, A., Sereshki, F., and Ataei, M. (2018). Development of over-break prediction models in drill and blast tunneling using soft computing methods. Engineering with computers. 34 (1): 45-58.

[22]. Khandelwal, M. and Singh, T.N. (2006). Prediction of blast induced ground vibrations and frequency in opencast mine: a neural network approach. Journal of sound and vibration. 289 (4-5): 711-725.

[23]. Rezaei, M., Monjezi, M., and Varjani, A.Y. (2011). Development of a fuzzy model to predict fly-rock in surface mining. Safety science. 49 (2): 298-305.

[24]. Lundborg, N. (1974). The hazards of fly-rock in rock blasting. Swedish Detonic Research Foundation, Reports DS, 12.

[25]. Hajihassani, M., Armaghani, D.J., Sohaei, H., Mohamad, E.T., and Marto, A. (2014). Prediction of airblast-overpressure induced by blasting using a hybrid artificial neural network and particle swarm optimization. Applied Acoustics. 80: 57-67.

[26]. Faradonbeh, R.S., Monjezi, M., and Armaghani, D.J. (2016). Genetic programing and non-linear multiple regression techniques to predict back-break in blasting operation. Engineering with computers. 32 (1): 123-133.

[27]. Ebrahimi, E., Monjezi, M., Khalesi, M.R., and Armaghani, D.J. (2016). Prediction and optimization of back-break and rock fragmentation using an artificial neural network and a bee colony algorithm. Bulletin of Engineering Geology and the Environment. 75 (1): 27-36.

[28]. Monjezi, M., Ahmadi, Z., Varjani, A.Y., and Khandelwal, M. (2013). Back-break prediction in the Chadormalu iron mine using artificial neural network. Neural Computing and Applications. 23 (3): 1101-1107.

[29]. Sayadi, A., Monjezi, M., Talebi, N., and Khandelwal, M. (2013). A comparative study on the application of various artificial neural networks to simultaneous prediction of rock fragmentation and back-break. Journal of Rock Mechanics and Geotechnical Engineering. 5 (4): 318-324.

[30]. Mohammadnejad, M., Gholami, R., Sereshki, F., and Jamshidi, A. (2013). A new methodology to predict backbreak in blasting operation. International journal of rock mechanics and mining sciences. (1997). 60: 75-81.

[31]. Esmaeili, M., Osanloo, M., Rashidinejad, F., Bazzazi, A.A., and Taji, M. (2014). Multiple regression, ANN and ANFIS models for prediction of back-break in the open-pit blasting. Engineering with computers. 30 (4): 549-558.

[32]. Monjezi, M., Bahrami, A., Varjani, A.Y., and Sayadi, A.R. (2011). Prediction and controlling of fly-rock in blasting operation using artificial neural network. Arabian Journal of Geosciences. 4 (3-4): 421-425.

[33]. Monjezi, M., Khoshalan, H.A., and Varjani, A.Y. (2012). Prediction of fly-rock and back-break in open-pit blasting operation: a neuro-genetic approach. Arabian Journal of Geosciences. 5 (3): 441-448.

[34]. Monjezi, M., Bahrami, A., and Yazdian Varjani, A. (2010). Simultaneous prediction of fragmentation and fly-rock in blasting operation using artificial neural networks. International journal of rock mechanics and mining sciences (1997). 47 (3): 476-480.

[35]. Monjezi, M. and Dehghani, H. (2008). Evaluation of effect of blasting pattern parameters on back-break using neural networks. International Journal of Rock Mechanics and Mining Sciences. 45 (8): 1446-1453.

[36]. Belkina, A.C., Ciccolella, C.O., Anno, R., Halpert, R., Spidlen, J., and Snyder-Cappione, J.E. (2019). Automated optimized parameters for T-distributed stochastic neighbor embedding improve visualization and analysis of large datasets. Nature communications. 10 (1): 1-12.

[37]. Zhu, W., Webb, Z. T., Mao, K., and Romagnoli, J. (2019). A deep learning approach for process data visualization using t-distributed stochastic neighbor embedding. Industrial and Engineering Chemistry Research. 58 (22): 9564-9575.

[38]. Hinton, G. and Roweis, S.T. (2002, December). Stochastic neighbor embedding. In NIPS (Vol. 15, pp. 833-840).

[39]. Van der Maaten, L. and Hinton, G. (2008). Visualizing data using t-SNE. Journal of machine learning research. 9 (11).

[40]. Rogovschi, N., Kitazono, J., Grozavu, N., Omori, T., and Ozawa, S. (2017, May). t-Distributed stochastic neighbor embedding spectral clustering. In 2017 International Joint Conference on Neural Networks (IJCNN) (pp. 1628-1632). IEEE.

[41]. Weinberger, K.Q., Sha, F., and Saul, L.K. (2004, July). Learning a kernel matrix for non-linear dimensionality reduction. In Proceedings of the twenty-first international conference on Machine learning (p. 106).

[42]. Tao, K., Cao, J., Wang, Y., Mi, J., Ma, W., and Shi, C. (2020). Chemometric Classification of Crude Oils in Complex Petroleum Systems using t-Distributed Stochastic Neighbor Embedding Machine Learning Algorithm. Energy and Fuels, 34(5), 5884-5899.

[43]. Liu, H., Yang, J., Ye, M., James, S.C., Tang, Z., Dong, J., and Xing, T. (2021). Using t-distributed Stochastic Neighbor Embedding (t-SNE) for cluster analysis and spatial zone delineation of groundwater geochemistry data. Journal of Hydrology, 597, 126146.

[44]. Arnø, M.L., Godhavn, J.M., and Aamo, O.M. (2021). At-bit estimation of rock density from real-time drilling data using deep learning with online calibration. Journal of Petroleum Science and Engineering, 109006.

[45]. Dorogush, A.V., Ershov, V., and Gulin, A. (2018). CatBoost: gradient boosting with categorical features support. arXiv preprint arXiv:1810.11363.

[46]. Prokhorenkova, L., Gusev, G., Vorobev, A., Dorogush, A.V., and Gulin, A. (2017). CatBoost: unbiased boosting with categorical features. arXiv preprint arXiv:1706.09516.

[47]. Huang, G., Wu, L., Ma, X., Zhang, W., Fan, J., Yu, X., ... and Zhou, H. (2019). Evaluation of CatBoost method for prediction of reference evapotranspiration in humid regions. Journal of Hydrology, 574, 1029-1041.

[48]. Hancock, J. and Khoshgoftaar, T.M. (2020, August). Medicare fraud detection using Catboost. In 2020 IEEE 21st international conference on information reuse and integration for data science (IRI) (pp. 97-103). IEEE.

[49]. Zhong, C., Geng, F., Zhang, X., Zhang, Z., Wu, Z., and Jiang, Y. (2021, May). Shear Wave Velocity Prediction of Carbonate Reservoirs Based on CatBoost. In 2021 4th International Conference on Artificial Intelligence and Big Data (ICAIBD) (pp. 622-626). IEEE.

[50]. Cardoso, A., Leitão, J., and Teixeira, C. (2018, September). Using the Jupyter notebook as a tool to support the teaching and learning processes in engineering courses. In International Conference on Interactive Collaborative Learning (pp. 227-236). Springer, Cham.

[51]. Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., ... and Duchesnay, E. (2011). Scikit-learn: Machine learning in Python. The Journal of machine Learning research, 12, 2825-2830.

[52] Althnian, A., AlSaeed, D., Al-Baity, H., Samha, A., Dris, A.B., Alzakari, N. and Kurdi, H. (2021). Impact of Dataset Size on Classification Performance: An Empirical Evaluation in the Medical Domain. Applied Sciences. 11 (2): 796.

[53]. Willmott, C.J. (1982). Some comments on the evaluation of model performance. Bulletin of the American Meteorological Society. 63 (11): 1309-1313.

[54]. Kamran, M. (2021). A Probabilistic Approach for Prediction of Drilling Rate Index using Ensemble Learning Technique. Journal of Mining and Environment.

[55]. Kamran, M., Bacha, S., and Mohammad, N. (2021). A Stochastic Model Updating Gold Reserve Estimation by Using Monte Carlo Simulation. Journal of Mining and Environment.

[56]. Guo, H., Nguyen, H., Bui, X.N., and Armaghani, D.J. (2021). A new technique to predict fly-rock in bench blasting based on an ensemble of support vector regression and GLMNET. Engineering with Computers. 37 (1): 421-435.

تکنیک توزیع جانشینی همسایه تصادفی مبتنی بر Catboost برای پیش‌بینی عقبزدگی در معدن سنگ آهک کارخانه سیمان دی‌وان، اندونزی

محمد کامران *

گروه مهندسی معدن، موسسه فناوری باندونگ، اندونزی

ارسال ۲۰۲۱/۰۹/۱۷، پذیرش ۲۰۲۱/۰۹/۲۷

* نویسنده مسئول مکاتبات: muhdkamran62@gmail.com

چکیده:

عملیات آتشکاری یک تکنیک مهم خردایش سنگ است که در چندین رشته مهندسی پایه مانند معدن، عمران، تونلسازی و جاده‌سازی استفاده می‌شود. عقبزدگی (BB) یکی از اثرات منفی ناشی از عملیات آتشکاری است که چندین اثر از جمله خسارت ماشین آلات معدنی، طراحی شیب پله و خطرات ناشی از الگوهای انفجارهای بعدی به دلیل خروج گازها از ناپیوستگی‌های ایجاد شده در توده سنگ را ایجاد می‌کند. چندین روش متعدد برای پیش‌بینی میزان عقبزدگی در عملیات آتشکاری توسط محققین مختلقی ارئه شده است. با این حال، این اولین کار برای پیاده‌سازی رویکرد پیشرفته مبتنی بر Catboost بر توزیع همسایه تصادفی (t-SNE) برای پیش‌بینی عقبزدگی است. در مجموع ۶۲ مجموعه داده دارای ۱۲ ویژگی تأثیرگذار در ایجاد عقبزدگی از الگوهای انفجار اصلی جمع آوری شده‌اند. یک روش جدید کاهش ابعاد t-SNE که تفسیر واگرایی Kullback-Leibler را اجرا می‌کند، برای تنظیم بزرگنمایی اولیه مجموعه داده انفجار استفاده شده است. سپس مجموعه داده t-SNE بدست آمده به نسبت ۷۰:۳۰ به مجموعه داده‌های آموزش و آزمایش تقسیم می‌شود. در نهایت، روش Cat-boost بر روی پایگاه داده انفجار با ابعاد کم اجرا می‌شود. معیار ارزیابی عملکرد تأیید می‌کند که مدل پیش‌بینی عقبزدگی از ثبات بیشتری برخوردار است و همپوشانی۹۹/۰۴ با مجموعه داده‌های آموزشی دارد و همچنین در مجموعه داده های آزمایش همپوشانی ۹۷/۲۶ را نشان داد. همچنین این مدل برای پیش‌بینی عقبزدگی عملکرد بهتری نسبت به مدل‌های مرسوم گذشته نشان داده است. به طور خلاصه ، این مدل را می توان به منظور پیش‌بینی عقبزدگی در چندین روش مهندسی سنگ و سناریوهای صنعت معدن به کار گرفت.

کلمات کلیدی: آتشکاری،عقبزدگی، Catboost، مهندسی سنگ، صنعت معدن.