Proc. of the 8th International Symposium on Geotechnical Safety and Risk (ISGSR)
Edited by Jinsong Huang, D.V. Griffiths, Shui-Hua Jiang, Anna Giacomini and Richard Kelly
©2022 ISGSR Organizers. Published by Research Publishing, Singapore.
doi: 10.3850/978-981-18-5182-7_08-013-cd

ISGSR 2022

Machine Learning-Based Prediction of Drilling and Blasting Tunnel Initial Support Patterns

Jinghan Chang[1], Hongwei Huang[1], Dongming Zhang[1], Tongjun Yang andZhenhua Xing[2]

[1]Department of Geotechnical Engineering, Tongji University, Shanghai, China.
E-mail: 2032347@tongji.edu.cn
[1]Department of Geotechnical Engineering, Tongji University, Shanghai, China.
E-mail: huanghw@mail.tongji.edu.cn
[1]Department of Geotechnical Engineering, Tongji University, Shanghai, China.
E-mail:09zhang@tongji.edu.cn
[2]Southwest Transportation Construction Group Co. Ltd, Kunming, Yunnan province, 650011, China
E-mail: 515066118@qq.com
[3]China state construction railway investment & engineering group CO.LTD, Beijing, China.
E-mail: 444840499@qq.com

Abstract: When the construction encounters a different geological condition from the design, the altered support patterns of China's NATM tunnel are designed by experienced experts based on the comprehensive evaluation of advanced geology forecast and their observation of the tunnel face. However, the current decision methods require expertise, and there are often subjective discrepancies. This study proposes a data-driven method for the support pattern decision. The inputs are thirteen variables describing the tunnel faces and the support pattern of the corresponding tunnel face recorded on the construction site are the outputs. We selected five classic machine learning classification algorithms (support vector machine(SVM), random forest, logistic regression, Gaussian process, and Naïve Bayes) through trial calculations and the optimal parameters of the models were obtained by using the grid search cross-validation technique. The results show that the voting classifier constructed with these five algorithms has good accuracy in the prediction of the support patterns. The feature importance rank of input variables is determined by sensitivity analysis, which enhances our understanding of the relationship between surrounding rock and support.

Keywords:Machine learning;rock tunnel;prediction;support patterns;tunnel face; decision support.

1 Introduction

With the rapid growth of traffic demand in China, the construction of mountain tunnels will continue to be an important research topic for Chinese civil builders and scholars. Among these constructions, the drilling and blasting method is still the most widely used, most important, and most commonly used mountain tunnel excavation method in China.

In the initial support design of drilling and blasting tunnels, the engineering analogy method is mainly used at present. When encountering dangerous sections, the numerical simulation will be used as an auxiliary decision-making method. In the actual tunnel excavation, due to the uncertainty of the surrounding rock, it is often the case that the support method given by the design does not match the conditions of the surrounding rock. Due to the convergent deformation characteristics of the surrounding rock, the construction of the initial support cannot be delayed. Based on this situation, we need a way to make support decisions as quickly as possible based on the surrounding rock conditions of the tunnel face on-site, while ensuring safety.

Before the advent of “big data” and advances in machine learning, it was appropriate to baserock engineering design primarily on empirical data and expert knowledge because the geology, discontinuity, and in situ stress data available for projects were generally sparse (Morgenroth 2021).Now, increasing amounts of some kinds of data such as Rock structure category, Groundwater area, Trace maps, are being collected relatively inexpensively by computer vision and deep learning(Chen et al. 2021).This is an opportunity to integrate machine learning into existing rock tunnel construction to evaluate rock data more effectively and extract value (or information) from the data to the maximum.

During the tunnel excavation, if support changes are required, the existing method is for experienced engineers to make decisions based on advanced drilling data and tunnel face conditions, and choose one of several designed support modes. This method is highly subjective and is largely limited by the experience level of the engineer himself, and it is easy to get a safer solution.

Aiming at the existing shortcomings of drilling and blasting tunnel construction, this paper constructs an initial support pattern prediction model based on the face data extracted by machine learning algorithms and computer vision. The database comes from published papers (Zhou 2021). Based on the statistics obtained from the Mengzi-Pingbian Highway Tunnel(MPHT) in Yunnan Province, China, five support patterns (i.e.,S4b, S4c, S5a, S5b, and S5c) were labeled based on the designing documents. We corresponded the 133 groups of tunnel faces in the database with the support patterns by consulting the construction record files. The input data of the model is composed of thirteen variables in the database, and the output is the support type. We processed the input data through a variety of data preprocessing methods, and then initially selected four machine learning models that performed well on the data. They are logistic regression(LR), random forest(RF), support vector machine(SVM), and AdaBoost Classifier(ABC). After optimizing the grid search parameters of these four models, the ABC with the best effect is selected.

This new method can help make quick decisions on-site, avoiding dangers such as collapse due to delayed timing, and also avoiding waste caused by improper support patterns.

2 Data pre-processing and model selection

In the design stage of the drill and blast tunnel, the designer will first preset several support patterns based on the geological survey data. Then choose different support modes for different sections. They will name the different support patterns S3, S4a, S4b, S5a, S5b, etc. The number represents the level of the surrounding rock, and the lowercase letters a, b, and c represent the different support patterns under the same level of the surrounding rock, and the support intensity increases sequentially. Figure 1shows the MPHT S5a Type Lining Design. Design parameters include bolt type, bolt arrangement, bolt length, shotcrete thickness, steel frame layout, etc.

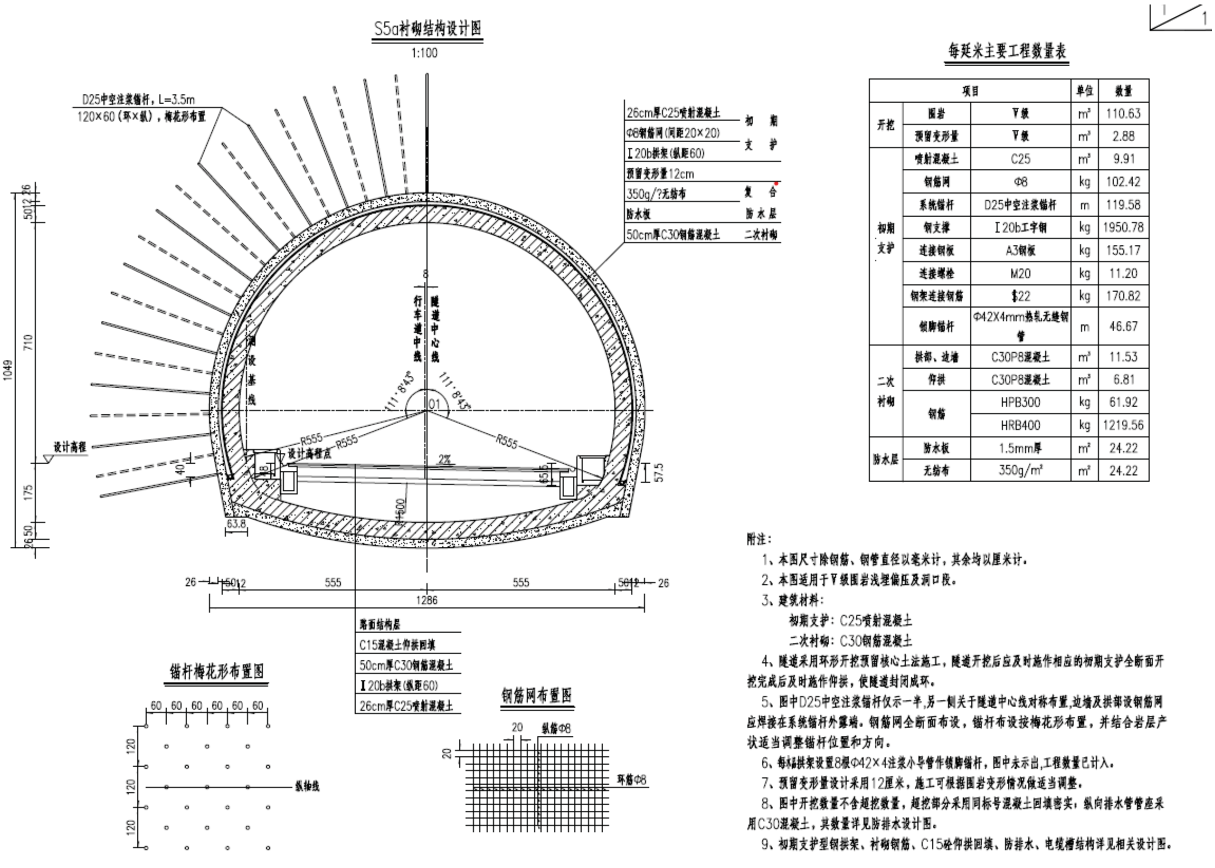

Figure 1. Mengzi-Pingbian Tunnel S5a Type Lining Design

Table 1reports the average, maximum as well as minimum values of the model outputs and inputs along with various other details. Among the thirteen features, ten are numerical features, and three are categorical features. Different features have different value ranges, and may even have different units. To obtain accurate classification results and ensure the function of each feature, the numerical features are standardized and the categorical features are numbered. This treatment can reduce the influence of model scale and dimension on the model. The number of each support pattern (S4b, S4c,S5a,S5b,S5c) in the sample is 30, 27, 20, 26, 30.Therefore, this data set does not have the problem of uneven samples.

Table 1. Summary of variable definitions and distribution

Variables	Symbol	Unit	Mean	Std	Min	Median	Max
Input							
Rock structure category	RSC	-	2.489	1.229	1	2	5
Groundwater category	GWC	-	2.406	1.200	1	2	5
Groundwater area	GWA	m^2	8.493	7.826	0	7.824	26.593
Weak interlayer area	WIA	m^2	3.521	2.752	0	3.051	10.330
Trace length (average)	TL_a	m	0.503	0.230	0.022	0.522	1.122
Trace length (maximum)	TL_{max}	m	0.922	0.258	0.283	0.928	1.637
Trace density	TD	-	40.309	23.346	1.046	36.605	103.540
Trace intensity	TI	-	7.801	6.383	0	6.275	25.101
Primary trace angle	PTA	°	97.705	50.636	3.378	113.309	186.849
Tunnel depth	H	m	221.624	87.352	68	217	370
Uniaxial compressive strength	UCS	MPa	20.084	12.247	3.55	17.50	54.89
Tunnel strike	TS	°	61.481	18.827	14.1	63.6	89.9
Weathering degree	WD	-	3.150	0.996	1	3	5
Output							
Support pattern	SP	-	2.992	1.490	1	3	5

Note:RSC categories: 1- block structure, 2- layered structure, 3- mosaic structure, 4- fragmentation structure, 5- granular structure; GWC categories: 1. dry state, 2. wet state, 3. dripping state, 4. flowing state, 5. gushing state; WD categories: 1. unweathered, 2. slightly weathered, 3. moderately weathered, 4. highly weathered, 5. decomposed. SP categories: 1-S4b, 2-S4c, 3-S5a, 4-S5b, 5-S5c.

We selected four more promising machine learning algorithms from eleven algorithms based on the accuracy of the model before tuning the parameters, namely: logistic regression (LR), random forest (RF), support vector machine (SVM), and AdaBoost. The dataset was randomly shuffled, of which 70% data were used as training and 30% for testing.Figure 2shows the accuracy of the 10-fold cross-validation of the eleven algorithms on the training set.

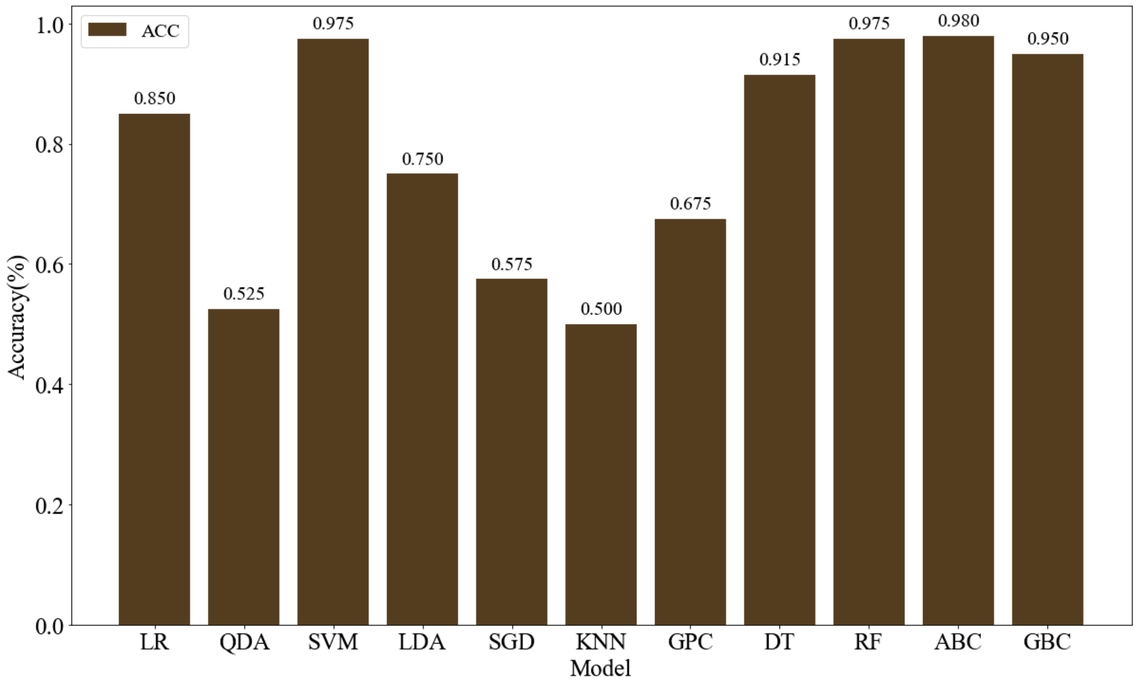

Figure 2. The accuracy of eleven algorithms

3 Support patterns prediction model

Based on the Scikit-learn machine learning basic algorithm package, this article uses the four supervised learning algorithms selected above to learn from the data training set, including conventional algorithms (LR and SVM) and integrated algorithms (RF and AdaBoost).In the process of model training, five-fold grid search cross-validation (GridSearchCV) is used to optimize the parameters of the model to obtain the optimal parameters when the accuracy of the algorithm is the highest. In the process of grid search, a series of prior candidate values of related parameters of the algorithm are first given, and all parameter value combinations are tried through loop traversal, and then the parameter value combinations that make the algorithm perform optimally are obtained. Table 2 presents the optimal parameters of each model through the grid search.

Table 2. Grid search and optimal parameters of four different models

Model	Parameter	Optional value	Optimal parameter	Time cost(s)
LR	*Penalty*	*['l1','l2']*	*'l1'*	10.752
	solver	*['newton-cg', 'lbfgs', 'liblinear', 'sag', 'saga']*	*'saga'*	
	Multi_class	*['ovr','multinomial']*	*'multinomial'*	
	C	*[0.01,0.05,0.1,0.25,0.5,0.75,1,10]*	*0.5*	
RF	*n estimators*	*[0,200]*	*40*	6424.713
	max depth	*[0,20]*	*15*	
	criterion	*'gini' or 'entropy'*	*'gini'*	
SVM	*Kernel*	*'linear', 'poly', 'rbf','sigmoid'*	*'linear'*	7.873
	C	*[0.05, 0.1, 0.25, 0.5, 0.75, 1.0]*	*1.0*	
AdaBoost	*n_estimators*	*range(1,100,5)*	*6*	
	algorithm	*['SAMME','SAMME.R']*	*'SAMME.R'*	186.455
	learning_rate	*np.arange(0.01,1,0.1)*	*0.71*	

Table 3shows the performances of the LR, RF, SVM and the AdaBoost methods on the training and test set. It can be seen that the AdaBoost has the best performance, but other methods only perform slightly worse. The average Precision, Recall and F-measure of all these methods on the training set are approximately 100%, while the average Precision, Recall and F-measure of these methods on the labeled test set are all above 90%.In terms of model calculation time, only the random forest algorithm has a slightly longer calculation time, and the other algorithms are all around 0.01 seconds.

Table 3. Performances of the proposed method and other supervised learning-based methods

Model	Training set (%)			Test set (%)			Time(s)
	P	*R*	F_1	*P*	*R*	F_1	
LR	99.00	99.00	98.97	90.56	91.14	90.61	0.012
RF	100.00	100.00	100.00	95.28	96.00	90.61	0.039
SVM	97.89	97.67	97.71	94.00	94.00	92.94	0.013
ABC	100.00	100.00	100.00	100.00	100.00	100.00	0.017

Figure 3shows the confusion matrixes of all four methods on the test set. The samples on the sub-diagonal of the matrix are correctly predicted. The bearing capacity of the surrounding rock above the sub-diagonal line is overestimated, and the model matches them with a weaker support pattern, which may lead to collapse; the bearing capacity of the surrounding rock below the sub-diagonal line is underestimated, The model matches them with a stronger support model, which will cause economic waste.

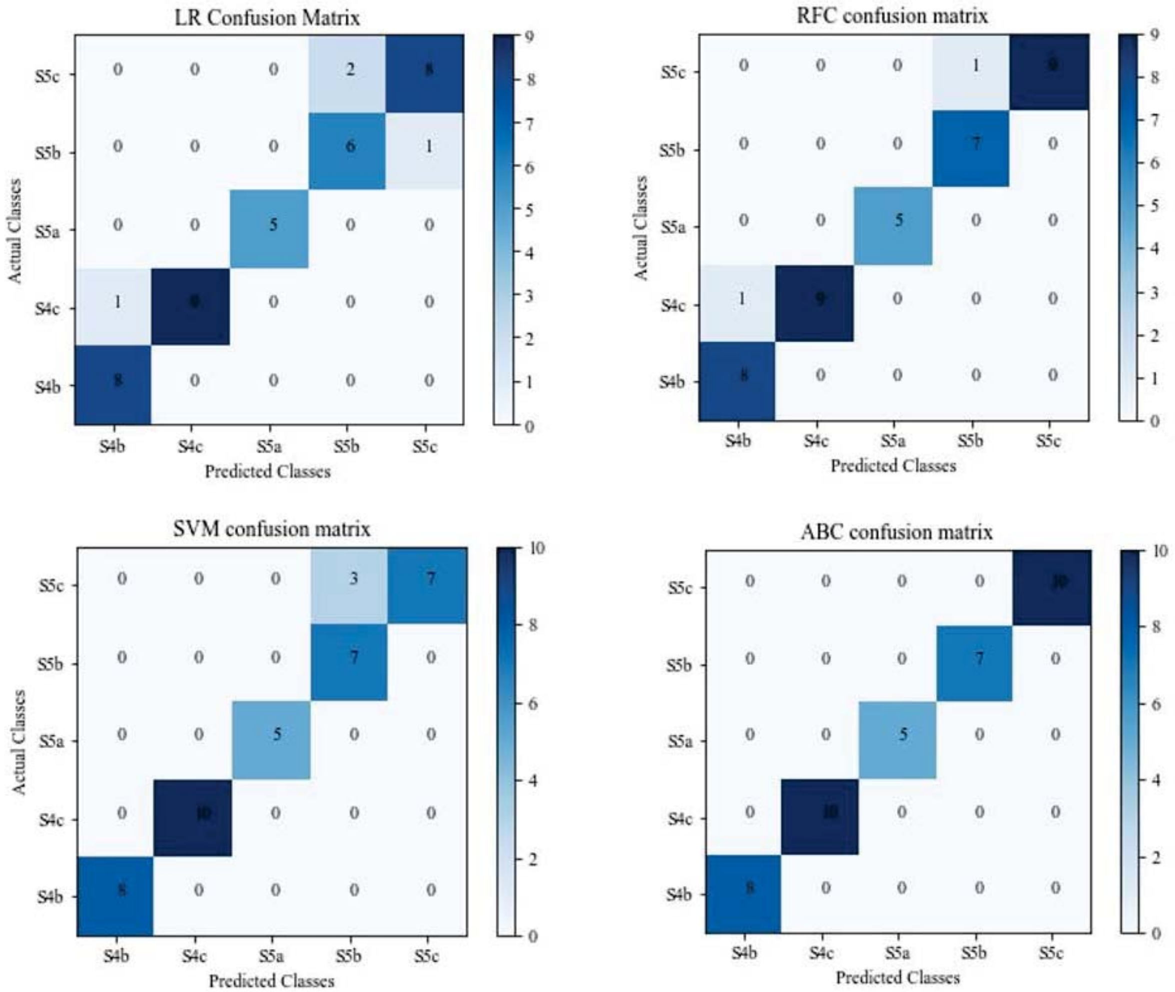

Figure 3. Confusion matrixes of four methods on the test set

4 Method discussion

Feature importance reflects the values of specific parameters in classification prediction. Therefore, it is necessary to calculate the importance of each feature to verify the rationality of the selected model input parameter.

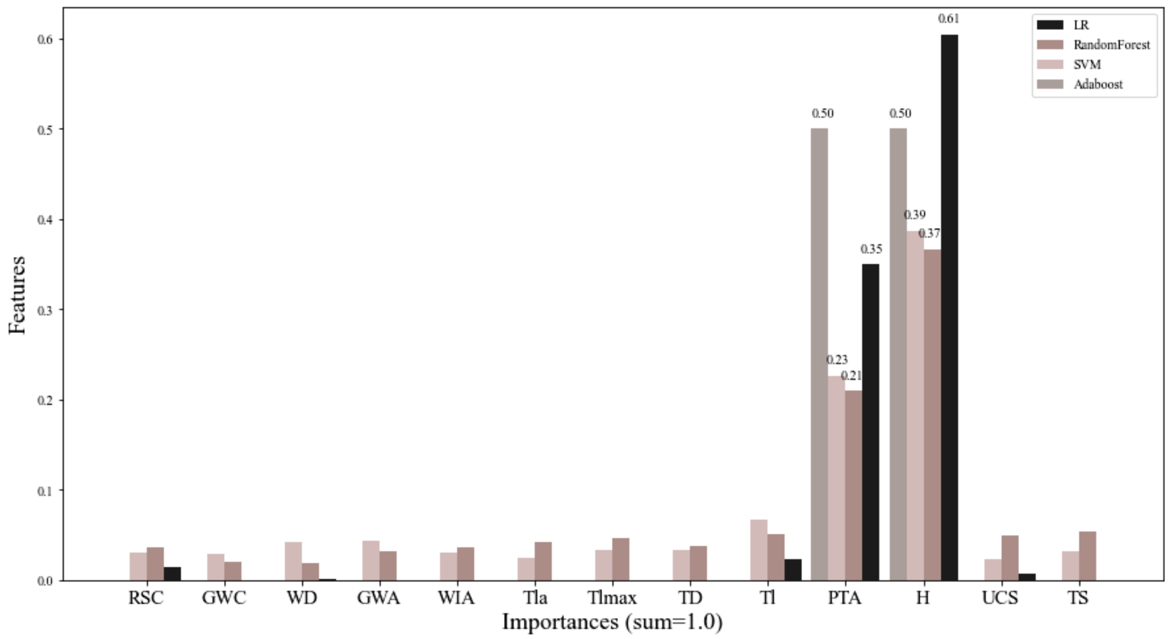

Figure 4. Importance scores of input features

In this study, the feature importance was evaluated by different method for different algorithms. For the random forest and AdaBoost model, the feature importance was evaluated by comparing the Gini index. The Gini index reflects the probability that two samples randomly selected from a dataset have inconsistent class labels, which can be calculated as

$$Gini = 1 - \sum_{k=1}^{K'} p_k^2 (1)$$

where K' is the number of support pattern classes, and p_k is the probabilitythat the sample belongs to the kth support pattern class.

For the SVM and LR model, the feature importance was evaluated by the weights assigned to the features. In order to unify the weight and the feature importance, the weight needs to be scaled to between 0 and 1.

Figure 4shows the importance scores of input features. For each model, the sum of the scores of the thirteen features is 1. Among these parameters, the tunnel depth (H) had the highest importance score followed by the primary trace angle (PTA), and other influencing factors are less important than the previous two.

5 Conclusions

In this study, a machine learning based support pattern prediction model was proposed. The tunnel face – support pattern database employed for the method was collected from the Mengping tunnel in Yunnan and 13 tunnel face parameters (RSC, GWC, WD, GWA, WIA, Tla, Tlmax, TD, Tl, PTA, H, UCS, TS) were used as the input features of the method. Overall, according to our research, the following conclusions can be drawn:

(1) In terms of prediction performance, the support pattern prediction model based on AdaBoost is significantly better than other mechine learning models. The Acc and F_1 values of the AdaBoost model are 1.00 and 1.00, higher than those of the LR, RF, SVM.

(2) The feature importance analysis shows that the contribution of each input features is different. The importance scores of input features show that, the tunnel depth (H) had the highest importance score followed by the primary trace angle (PTA) , and other influencing factors are less important than the previous two.

References

Morgenroth, J., Khan, U. T., & Perras, M. A. (2019). An overview of opportunities for machine learning methods in underground rock engineering design. *Geosciences*, 9(12), 504.

Chen, J., Yang, T., Zhang, D., Huang, H., & Tian, Y. (2021). Deep learning based classification of rock structure of tunnel face. *Geoscience Frontiers*, 12(1), 395-404.

Chen, J., Zhou, M., Zhang, D., Huang, H., & Zhang, F. (2021). Quantification of water inflow in rock tunnel faces via convolutional neural network approach. *Automation in Construction*, 123, 103526.

Chen, J., Zhou, M., Huang, H., Zhang, D., & Peng, Z. (2021). Automated extraction and evaluation of fracture trace maps from rock tunnel face images via deep learning. *International Journal of Rock Mechanics and Mining Sciences*, 142, 104745.

Zhou, M., Chen, J., Huang, H., Zhang, D., Zhao, S., & Shadabfar, M. (2021). Multi-source data drive method for assessing the rock mass quality of a NATM tunnel face via hybrid ensemble learning models. *International Journal of Rock Mechanics and Mining Sciences*, 147, 104914.