



ELSEVIER

Contents lists available at ScienceDirect

MethodsX

journal homepage: www.elsevier.com/locate/mex

Method Article

Method for geological characteristics prediction during shield tunnelling: SCA-GS[☆]

Tao Yan^{a,b,*}

^aMOE Key Laboratory of Intelligent Manufacturing Technology, Department of Civil and Environmental Engineering, College of Engineering, Shantou University, Shantou, Guangdong, 515063, China

^bDiscipline of Civil and Infrastructure, School of Engineering, Royal Melbourne Institute of Technology (RMIT), Victoria, 3001, Australia

A B S T R A C T

Geological characteristic (GC) is one of the most essential factors influencing setting earth pressure balance (EPB) shield parameters and cutterhead wear. Identification of GC has crucial significance to shield tunnelling efficiency and safety. Stacking classification algorithm (SCA) is widely applied in engineering with the identification and classification. Grid search (GS) is designed to tune hyper-parameter and optimize non-linear problems with K-folds cross-validation (K-CV), which is commonly used to change validation set in the training set. The performance of SCA can be improved by GS and K-CV. The types of GC during shield advance can be identified by integrating K-means++ with silhouette coefficient (S_i) and elbow method (EM). The results of K-means++ and shield parameters severed as a database for SCA. The approach was applied in Guangzhou mixed ground. The results showed that the proposed framework could predict the geological characteristics well. The method article is a companion paper with the original article [1]. The proposed method enables:

- Developed approach merges SCA and GS method.
- Application of SCA-GS method in geological characteristics classification.
- It can increase the reliability of classification results.

© 2022 The Author(s). Published by Elsevier B.V.

This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>)

A R T I C L E I N F O

Method name: Integrating SCA method with GS and K-CV

Keywords: Stacking classification algorithm, Grid search, K-folds cross-validation, prediction

Article history: Received 21 April 2022; Accepted 10 October 2022; Available online 20 October 2022

[☆] ARTICLE INFO Method name: Method of integrating GS and K-CV into stacking classification algorithm

* Corresponding author: College of Engineering, Shantou University, Shantou, Guangdong 515063, China

E-mail address: yantao@stu.edu.cn

Specification Table

Subject area:	Engineering Geology
More specific subject area:	Geological characteristics
Method:	Integrating SCA method with GS and K-CV
Name and reference of the original method:	Wolpert, D.H., (1992) Stacked generalization, Neural Networks, 5(2), 241-259, https://doi.org/10.1016/S0893-6080(05)80023-1 [2].
Resource availability:	DOI: https://doi.org/10.1016/j.jrmge.2022.03.002 .

Method details

Stacking classification algorithm (SCA)

In the process of SCA, a variety of models were combined into an ensemble algorithm. The detail of SCA implementation is given [2]:

Step 1: Establish the database and pre-process data. The data with different characteristics are acquired, and the database matrix is constructed in Eq. (1).

$$X = [x_{i,j}]_{n \times m} = \begin{bmatrix} & C_1 & C_2 & C_3 & \dots & C_m \\ X_1 & x_{1,1} & x_{1,2} & x_{1,3} & \dots & x_{1,m} \\ X_2 & x_{2,1} & x_{2,2} & x_{2,3} & \dots & x_{2,m} \\ X_3 & x_{3,1} & x_{3,2} & x_{3,3} & \dots & x_{3,m} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ X_n & x_{n,1} & x_{n,2} & x_{n,3} & \dots & x_{n,m} \end{bmatrix} \tag{1}$$

where C_m is the name of data m ; X_n denotes the n^{th} sample in C_m . Then, the mean value of 3 sequential points after time t is used to replace the value of time t . Meanwhile, the data in X_n are normalized to the interval [0, 1] according to Min-Max Normalization. The normalized database is presented as follow:

$$X' = [x'_{i,j}]_{n \times m} = \begin{bmatrix} & C_1 & C_2 & C_3 & \dots & C_m \\ X'_1 & x'_{1,1} & x'_{1,2} & x'_{1,3} & \dots & x'_{1,m} \\ X'_2 & x'_{2,1} & x'_{2,2} & x'_{2,3} & \dots & x'_{2,m} \\ X'_3 & x'_{3,1} & x'_{3,2} & x'_{3,3} & \dots & x'_{3,m} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ X'_n & x'_{n,1} & x'_{n,2} & x'_{n,3} & \dots & x'_{n,m} \end{bmatrix} \tag{2}$$

where X' and x' are the normalized value corresponding to Eq.(1).

Step 2: Determine types for the database. The database types can be labelled by K-means++ algorithm [3]. Euclidean distance is applied in K-means++, which can be expressed as Eq. (3).

$$dist(A, B) = \sqrt{\sum_{i=1}^n |a_i - b_i|^2} \tag{3}$$

where $dist(A, B)$ represents the distance of samples A and B; a_i and b_i denote the coordinates of samples A and B. The square error is utilized as the objective function in K-means++. The objective function is a convex function of cluster centers ($\mu_1, \mu_2, \dots, \mu_k$) and the stagnation point of the objective function is the clustering centers. The objective function and cluster centers can be expressed as follows [4]:

$$J(\mu_1, \mu_2, \dots, \mu_k) = \frac{1}{2} \sum_{j=1}^K \sum_{i=1}^{N_j} (x_i - \mu_j)^2 \tag{4}$$

$$\frac{\partial J}{\partial \mu_j} = -2 \sum_{i=1}^{N_j} (x_j - \mu_j) \rightarrow 0 \Rightarrow \mu_j = \frac{1}{N_j} \sum_{i=1}^{N_j} x_j \tag{5}$$

where $J()$ is the objective function; μ_j is the center of cluster j ; N_j denotes the sample number of category j ; x_i represents a random sample in the database; x_j is the sample point of cluster j . The elbow method (EM) and silhouette coefficient (S_i) are utilized to find the value of K in K -means++. The diagram of the sum of the squared errors (SSE) versus K in elbow method will form an elbow, and the corresponding K to the elbow of the diagram is determined as the optimal cluster number. When there are more than one elbow in elbow method, K in elbow method corresponding to the maximum silhouette coefficient (S_i) can be selected as the best value for the K -means++. The SSE and silhouette coefficient (S_i) can be calculated by Eq. (4) and Eq. (5) in the companion paper [1]. The value of K types can be set as 1 to K corresponding to different values of X_n . Then, each X_n will be labelled with the value of K (L_k^n) in the matrix in Eq. (6).

$$X'_k = [x'_{i,j}, L]_{n \times (m+1)} = \begin{bmatrix} & C_1 & C_2 & C_3 & \dots & C_m & Label \\ X'_1 & x'_{1,1} & x'_{1,2} & x'_{1,3} & \dots & x'_{1,m} & L_{k,1}^1 \\ X'_2 & x'_{2,1} & x'_{2,2} & x'_{2,3} & \dots & x'_{2,m} & L_{k,2}^2 \\ X'_3 & x'_{3,1} & x'_{3,2} & x'_{3,3} & \dots & x'_{3,m} & L_{k,3}^3 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ X'_n & x'_{n,1} & x'_{n,2} & x'_{n,3} & \dots & x'_{n,m} & L_{k,n}^n \end{bmatrix} \quad (6)$$

Step 3: Input primary learners, meta-classifier and establish stacking classification model. SCA includes two layers, primary learners and meta-classifier. Support vector machine (SVM), random forest (RF), and gradient boosting decision tree (GBDT) are utilized in the primary learners [5]. The logistic regression algorithm (LR) is employed as a meta-classifier. The results of the first layer with more characteristics can be taken as the input in the second layer to obtain higher accuracy for the proposed model. The results of the first layer can be constructed in Eq. (7).

$$X'_{k,3} = [x'_{i,j}, L]_{n \times (m+3)} = \begin{bmatrix} & C_1 & C_2 & C_3 & \dots & C_m & Label_1 & Label_2 & Label_3 \\ X'_1 & x'_{1,1} & x'_{1,2} & x'_{1,3} & \dots & x'_{1,m} & L_{k,1}^1 & L_{k,2}^1 & L_{k,3}^1 \\ X'_2 & x'_{2,1} & x'_{2,2} & x'_{2,3} & \dots & x'_{2,m} & L_{k,1}^2 & L_{k,2}^2 & L_{k,3}^2 \\ X'_3 & x'_{3,1} & x'_{3,2} & x'_{3,3} & \dots & x'_{3,m} & L_{k,1}^3 & L_{k,2}^3 & L_{k,3}^3 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ X'_n & x'_{n,1} & x'_{n,2} & x'_{n,3} & \dots & x'_{n,m} & L_{k,1}^n & L_{k,2}^n & L_{k,3}^n \end{bmatrix} \quad (7)$$

where $L_{k,m}^n$ is the label of sample n in model m ($m = 1, 2, 3$). The radial basis function (rbf) in SVM, the method of weighted mean in RF, the multi-classification loss function in GBDT are expressed as Eq. (8)-(10) and applied to classify the multiple geological characteristics [6-8].

$$\kappa(x_i, x_j) = \exp\left(-\frac{\|x_i - x_j\|^2}{2\sigma^2}\right) \quad (8)$$

$$G(x) = \sum_{i=1}^m w_i g_i(x) \quad (9)$$

$$L(y, f_i(x)) = -\sum_{k=1}^K y_k \log p_k(x), y_k \in (-1, 1) \quad (10)$$

where $\kappa()$ is the kernel function; x_i, x_j are sample points; $G(x)$ is the output according to weight of x ; w_i is the weight of decision trees; $g_i(x)$ is the value of decision trees; $L()$ represents the loss function; $f_i(x)$ is the output function; y denotes the label of point x ; $p_k(x)$ represents the probability for sample x in class k .

Step 4: Train the stacking classification model and predict. The classification model will be trained and evaluated using cross-validation and training set. Fig. 1 indicates the frame of the first layer running process in stacking algorithm (taking 3-CV as a case). The training set is input into the first layer to train the primary learners and establish the primary models (M_i). The new features in training set and test set will be calculated via the primary models using the validation set in raw training and

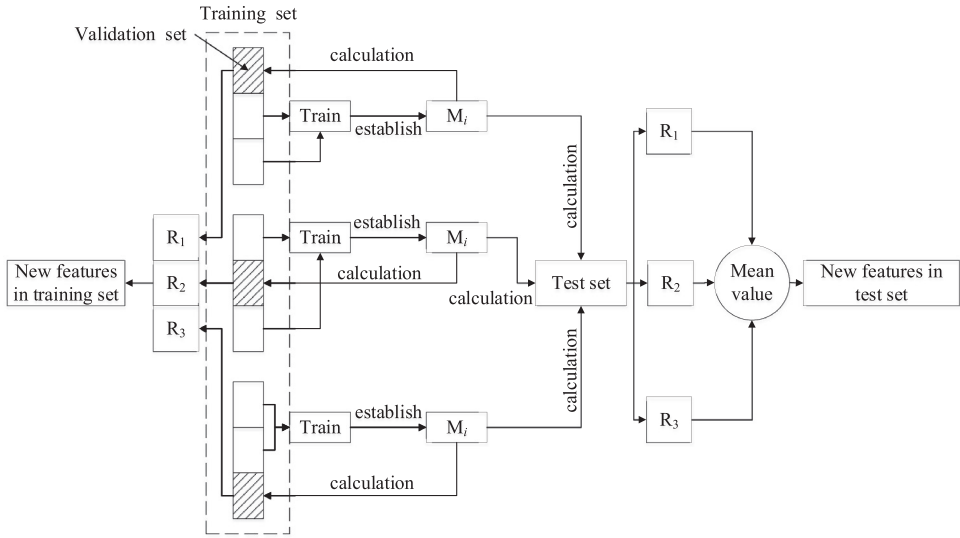


Fig. 1. First layer running process in stacking classification algorithm (SCA).

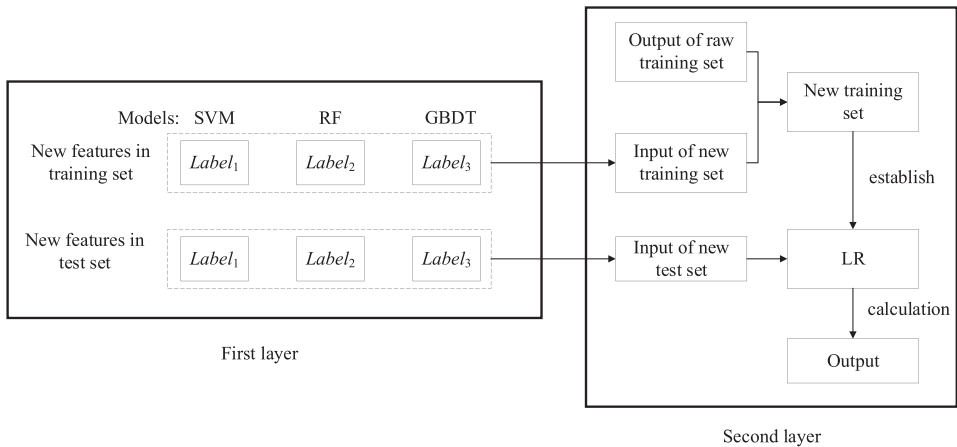


Fig. 2. Second layer running process in stacking classification algorithm (SCA).

test sets. Fig. 2 shows the second layer running process in stacking classification algorithm. The new features in training set are taken as input of new training set. The new training set is composed of input of new training set and output of raw training set, which are applied to train the classifier in second layer (LR) and establish a prediction model. Then, the new features in test set are taken as input of new test set and are input into the prediction model to output the final results.

Step 5: Evaluate the stacking classification algorithm model. The output GC of the SCA algorithm will be compared with the actual geological characteristics of raw test set. The F1-score, recall (R), and precision (P) are used to assess the proposed strategy.

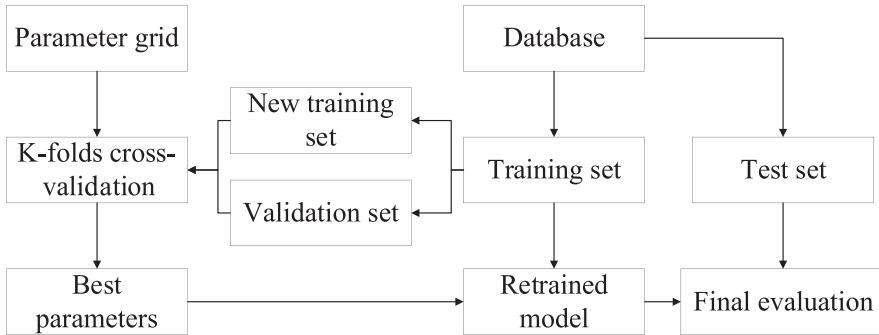


Fig. 3. Flowchart of grid search and K-folds cross-validation.

Grid search with cross-validation

Grid search is an exhaustive search method of selecting parameters. In all candidate parameters, every possibility will be tried, and the best performance of parameter combination will be selected as final parameters of the model. However, the database is divided into test and training set, and the model with the default hyper-parameters are trained using training set, which has low accuracy. The grid search combined with cross-validation can solve the problem of high error for the model. The database is divided into three parts: test, training, and validation set. The proposed framework can be fitted by training set and verified on validation set. Test set is applied in assessing the model. To avoid the influence of database division, cross-validation is employed to reduce contingency of database division. K-folds cross-validation is commonly used to change validation set in the training set. Grid search is usually applied together with K-CV. Fig. 3 shows the flowchart of GS and K-CV.

Application of the method

The SCA-GS model can be used to classify the GC during shield advance. Four shield parameters (cutterhead rotation speed (CRS), advance rate (AR), shield thrust (F), and cutterhead torque (T)) were collected from the sensors installed in EPB shield machine [9–11]. The empty data is removed according to Eq. (11) and Eq. (12) [12].

$$D = f(F) \cdot f(AD) \cdot f(CRS) \cdot f(T) \tag{11}$$

$$f(x) = \begin{cases} 0, & x = 0 \\ 1, & x \neq 0 \end{cases} \tag{12}$$

where if $D = 1$, data was saved. If $D = 0$, data was excluded. x is the value of shield parameters. The abnormal value out of three times standard deviation from the mean value was removed with Pauta criterion. The original data matrix is constructed in Eq. (13).

$$X = [x_{i,j}]_{n \times 4} = \begin{bmatrix} & F & AD & CRS & T \\ X_1 & 10.5 & 21 & 1.1 & 2.5 \\ X_2 & 27.5 & 15 & 1.2 & 5.7 \\ X_3 & 27 & 10 & 1.8 & 3 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ X_{1179} & 18.7 & 35 & 1.3 & 2.3 \end{bmatrix} \tag{13}$$

where F is shield thrust ($\times 10^3$ kN); AR is advance rate (mm/min); CRS is cutterhead rotation speed (rpm); T denotes cutterhead torque ($\times 10^3$ kN·m). Then, the raw data were transformed into torque penetration index (TPI) and field penetration index (FPI) [13, 14], which can be obtained by Eq. (6) in

the associated paper [1]. The mean value of 3 sequential points after time t replaced the value of time t . Meanwhile, TPI and FPI are normalized to the interval [0, 1] according to Min-Max Normalization. The normalized matrix is taken as input data and constructed in Eq. (14).

$$X' = [x'_{i,j}]_{n \times 2} = \begin{bmatrix} & FPI & TPI \\ X'_1 & 0.279 & 0.014 \\ X'_2 & 0.120 & 0.142 \\ X'_3 & 0.289 & 0.172 \\ \vdots & \vdots & \vdots \\ X'_{1179} & 0.023 & 0.012 \end{bmatrix} \tag{14}$$

Geological characteristics are labelled by K-means++ algorithm with elbow method (EM) and silhouette coefficient (S_j) using normalized FPI and TPI. Then, according to the site investigation of the construction site, the GCs in the project can be identified as K types, and the value of K types can be set as 1 to K, corresponding to different values of FPI and TPI, respectively. The geological characteristics in each tunnelling ring were labelled with the value of K (L'_n) in Fig. 8 in the companion paper [1], and the database matrix is constructed in Eq. (15).

$$X'_k = [x'_{i,j}, L]_{n \times 3} = \begin{bmatrix} & FPI & TPI & Label \\ X'_1 & 0.279 & 0.014 & 1 \\ X'_2 & 0.120 & 0.142 & 2 \\ X'_3 & 0.289 & 0.172 & 3 \\ \vdots & \vdots & \vdots & \vdots \\ X'_{1179} & 0.023 & 0.012 & 1 \end{bmatrix} \tag{15}$$

The database in Eq. (15) was split into training set with 80% and test set with 20% of data set and were input to the SCA-GS model. The optimal value of parameters and accuracy of primary learners are given in Table 3 in the companion paper [1]. The algorithms in the first layer with optimal hyperparameters are trained and tested using training and test sets. The results of the first layer can be constructed in Eq. (16).

$$X'_{k3} = [x'_{i,j}, L]_{n \times 5} = \begin{bmatrix} & FPI & TPI & Label_1 & Label_2 & Label_3 \\ X'_1 & 0.279 & 0.014 & 1 & 1 & 1 \\ X'_2 & 0.120 & 0.142 & 2 & 2 & 2 \\ X'_3 & 0.289 & 0.172 & 3 & 3 & 3 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ X'_{1179} & 0.023 & 0.012 & 1 & 1 & 1 \end{bmatrix} \tag{16}$$

The final output of prediction model was given in Fig. 13 in the companion paper [1]. The performance of SCA-GS prediction model was improved with the highest accuracy (0.996), which satisfies the requirement of shield tunnelling. Therefore, the prediction model can be utilized in a new project.

Computational tool

The study used Python program to establish the SCA-GS prediction model for geological characteristics during shield tunnelling. The pseudocode of improved stacking classification algorithm is listed in the Appendix. The source code of the SCA-GS includes cluster, optimization and prediction modules. Fig. 4 shows the flowchart of geological characteristics prediction. The detailed steps of the method application are presented as follows:

- (1) The users should prepare the dataset, which consists of the historical shield parameters (F, AR, CRS, and T) collected from shield operational system. Then, the empty data were removed based on Eq. (11) and (12).

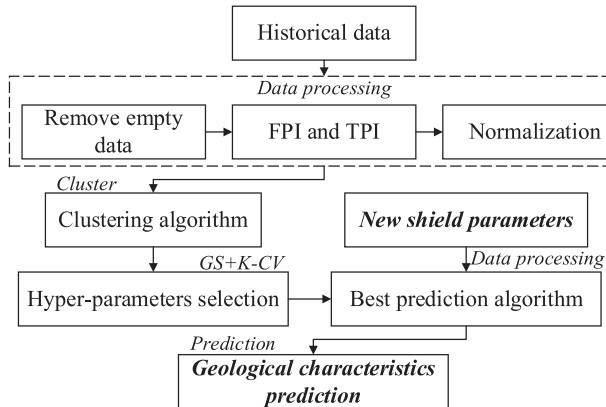


Fig. 4. Flowchart of geological characteristics prediction.

- (2) The shield parameters were calculated as FPI and TPI, which were normalized to the interval [0, 1] and input into the first part of the source code (cluster module). The users can change the data path of source code to input them.
- (3) The users can select the first cluster module to run the K-means ++ algorithm to give each line a label (GC). The results of the cluster module were saved as the data set for the following steps.
- (4) The data path was changed as the path of the results of K-means ++. The data set was input into the optimization module, which was used for training and testing prediction algorithms. The optimizer (GS and K-CV) will optimize the primary learners and provide the best hyper-parameters for prediction module.
- (5) The hyper-parameters of the prediction module were replaced based on the results of step (4). As such, the best prediction model was established with the highest accuracy of prediction algorithms.
- (6) The new shield parameters can be obtained during shield advance. Then, the users can set the input data path as the processed shield parameters and run the best prediction model to forecast real-time geological characteristics

Readers can contact the author to apply for the source code.

Method validation

To verify the advantage of the designed SCA-GS method, the existing SVM, RF, GBDT, and SCA-GS techniques were used to recognize the relationship between GC and shield parameters in Fig. 13 in the companion paper [1]. Based on the accuracy and performance of four classification algorithms, the SCA-GS is the better prediction model for geological characteristics. The detailed comparison and analyses of prediction results can be found in the companioned research article [1].

Identifying geological characteristics is crucial for shield tunnelling and reducing construction risk [15–17]. However, there is no globally accepted mechanical relationship between shield parameters and geological characteristics. The variation of shield parameters is a gray-box process based on geological features. Many factors, e.g., ground settlement, underground water, and lining quality, may influence the tunnelling process [18–20]. Artificial intelligence, including expert systems, machine learning, and deep learning, is an excellent technique for establishing the relationship between various parameters and objectives [21–24]. Four shield parameters were selected in this study to classify GC during shield advance. Besides, more parameters, e.g., specific energy, cutter wear, and earth pressure, should also be considered to evaluate their contribution to the prediction of GC [25].

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

CRedit authorship contribution statement

Tao Yan: Data curation, Methodology, Investigation, Software, Writing – original draft, Writing – review & editing.

Acknowledgments

The author would like to thank Shantou University and Royal Melbourne Institute of Technology (RMIT) for technical support. In addition, the author would like to thank reviewers for their comments to improve this article.

Appendix

Table A1.

Table A1

Pseudocode of SCA-GS.

Algorithm SCA-GS classification algorithm

```

1: input: training data  $A = \{(a_1, b_1), (a_2, b_2), \dots, (a_m, b_m)\}$ ;
primary learner:  $M_1, M_2, \dots, M_S$ ;
meta-classifier:  $M$ .
2. for  $s = 1, 2, \dots, S$  do
   $Param_s = \{p_1: (a_1, b_1, t_1), p_2: (a_2, b_2, t_2), \dots, p_n: (a_n, b_n, t_n)\}$ 
   $GS_s = \text{grid search}(M_s, Param_s, K\text{-CV})$ 
3:for  $s = 1, 2, \dots, S$  do
  for  $j = 1, 2, \dots, m$  do
     $M_s' = \text{fit}(GS_s(a_j))$ 
     $M_s'' = \text{high\_accuracy}(M_s')$ 
  end for
end for
4. for  $s = 1, 2, \dots, S$  do
   $r_s = M_s''(A)$ ;
   $A' = \emptyset$ ;
end for
5. for  $j = 1, 2, \dots, m$  do
  for  $s = 1, 2, \dots, S$  do
     $u_{js} = r_s(a_j)$ ;
  end for
   $A' = A' \cup \{(u_{j1}, u_{j2}, \dots, u_{jS}), b_j\}$ 
end for
6.  $r' = M(A')$ 
7. Output:  $R(a) = r'(r_1(a), r_2(a), \dots, r_S(a))$ 

```

References

- [1] T. Yan, S.L. Shen, A. Zhou, Prediction of geological characteristics from shield operational parameters by integrating grid search and K-fold cross validation into stacking classification algorithm, *J. Rock Mech. Geotech. Eng.* 14 (2022) 1292–1303, doi:[10.1016/j.jrmge.2022.03.002](https://doi.org/10.1016/j.jrmge.2022.03.002).
- [2] D.H. Wolpert, Stacked generalization, *Neural Netw.* 5 (2) (1992) 241–259, doi:[10.1016/S0893-6080\(05\)80023-1](https://doi.org/10.1016/S0893-6080(05)80023-1).
- [3] K. Elbaz, T. Yan, A. Zhou, S.L. Shen, Deep learning analysis for energy consumption of shield tunneling machine drive system, *Tunnell. Underground Space Technol.* 123 (2022) 104405, doi:[10.1016/j.tust.2022.104405](https://doi.org/10.1016/j.tust.2022.104405).
- [4] R.T. Aldahdooh, W. Ashour, DIMK-means "Distance-based Initialization Method for K-means Clustering Algorithm", *Int. J. Intell. Syst. Appl.* 5 (2074-904X) (2013) 41–51, doi:[10.5815/ijisa.2013.02.05](https://doi.org/10.5815/ijisa.2013.02.05).

- [5] P. Zhang, Y.F. Jin, Z.Y. Yin, Y. Yang, Random forest based artificial intelligent model for predicting failure envelopes of caisson foundations in sand, *Appl. Ocean Res.* 101 (2020) 102223, doi:[10.1016/j.apor.2020.102223](https://doi.org/10.1016/j.apor.2020.102223).
- [6] B. Scholkopf, K.K. Sung, C.J.C. Burges, F. Girosi, P. Niyogi, T. Poggio, V. Vapnik, Comparing support vector machines with Gaussian kernels to radial basis function classifiers, *IEEE Trans. Signal Process.* 45 (11) (1997) 2758–2765, doi:[10.1109/78.650102](https://doi.org/10.1109/78.650102).
- [7] V.F. Rodriguez-Galiano, B. Ghimire, J. Rogan, M. Chica-Olmo, J.P. Rigol-Sanchez, An assessment of the effectiveness of a random forest classifier for land-cover classification, *ISPRS J. Photogramm. Remote Sens.* 67 (2012) 93–104, doi:[10.1016/j.isprsjprs.2011.11.002](https://doi.org/10.1016/j.isprsjprs.2011.11.002).
- [8] J.H. Friedman, Greedy function approximation: a gradient boosting machine, *Ann. Stat.* 29 (5) (2000), doi:[10.1214/aos/1013203451](https://doi.org/10.1214/aos/1013203451).
- [9] S.L. Shen, K. Elbaz, W.M. Shaban, A. Zhou, Real-time prediction of shield moving trajectory during tunnelling, *Acta Geotech.* 17 (2022) 1533–1549, doi:[10.1007/s11440-022-01461-4](https://doi.org/10.1007/s11440-022-01461-4).
- [10] H.M. Lyu, S.L. Shen, A. Zhou, Z.Y. Yin, Assessment of safety status of shield tunnelling using operational parameters with enhanced SPA, *Tunnell. Underground Space Technol.* 123 (2022) 104428, doi:[10.1016/j.tust.2022.104428](https://doi.org/10.1016/j.tust.2022.104428).
- [11] S.S. Lin, N. Zhang, A. Zhou, Time-series prediction of shield movement performance during tunneling based on hybrid model, *Tunnell. Underground Space Technol.* 119 (2022) 104245, doi:[10.1016/j.tust.2021.104245](https://doi.org/10.1016/j.tust.2021.104245).
- [12] Q.L. Zhang, Z.Y. Liu, J.R. Tan, Prediction of geological conditions for a tunnel boring machine using big operational data, *Autom. Constr.* 100 (2019) 73–83, doi:[10.1016/j.autcon.2018.12.022](https://doi.org/10.1016/j.autcon.2018.12.022).
- [13] P.J. Tarkoy, M. Marconi, Difficult rock comminution and associated geological conditions, in: *Tunnelling' 91, Sixth International Symposium, Institute of Mining and Metallurgy, London, 1991*, pp. 195–207.
- [14] T. Yan, S.L. Shen, A. Zhou, H.M. Lyu, Construction efficiency of shield tunnelling through soft deposit in Tianjin, China, *Tunnell. Underground Space Technol.* 112 (2021) 103917, doi:[10.1016/j.tust.2021.103917](https://doi.org/10.1016/j.tust.2021.103917).
- [15] Q. Zheng, S.L. Shen, A. Zhou, H.M. Lyu, Inundation risk assessment based on G-DEMATEL-AHP and its application to Zhengzhou flooding disaster, *Sustain. Cities Soc.* 86 (2022) 104138, doi:[10.1016/j.scs.2022.104138](https://doi.org/10.1016/j.scs.2022.104138).
- [16] H.M. Lyu, W.J. Sun, A. Zhou, Risk assessment using a new consulting process in fuzzy, AHP, *J. Construct. Eng. Manag.* 146 (3) (2020) 04019112, doi:[10.1061/\(ASCE\)CO.1943-7862.0001757](https://doi.org/10.1061/(ASCE)CO.1943-7862.0001757).
- [17] Y.L. Chen, S.L. Shen, A. Zhou, Assessment of red tide risk by integrating CRITIC weight method, TOPSIS-ASSETS method, and Monte Carlo simulation, *Environ. Pollut.* 310 (2022) 120254, doi:[10.1016/j.envpol.2022.120254](https://doi.org/10.1016/j.envpol.2022.120254).
- [18] H.N. Wu, S.L. Shen, R.P. Chen, A. Zhou, Three-dimensional numerical modelling on localised leakage in segmental lining of shield tunnels, *Comput. Geotech.* 122 (2020) 10354, doi:[10.1016/j.compgeo.2020.103549](https://doi.org/10.1016/j.compgeo.2020.103549).
- [19] P. Zhang, Y.F. Jin, Z.Y. Yin, Machine learning-based uncertainty modeling of mechanical properties of soft clays relating to time-dependent behavior and its application, *Int. J. Numer. Anal. Methods Geomech.* 45 (11) (2021) 1588–1602, doi:[10.1002/nag.3215](https://doi.org/10.1002/nag.3215).
- [20] J.X. Zhang, N. Zhang, A. Zhou, S.L. Shen, Numerical evaluation of segmental tunnel lining with voids in outside backfill, *Underground Space* 7 (5) (2022) 786–797, doi:[10.1016/j.undsp.2021.12.007](https://doi.org/10.1016/j.undsp.2021.12.007).
- [21] N. Zhang, S.L. Shen, A. Zhou, Y.F. Jin, Application of LSTM approach for modelling stress-strain behavior of soil, *Appl. Soft Comput.* 100 (2021) 106959, doi:[10.1016/j.asoc.2020.106959](https://doi.org/10.1016/j.asoc.2020.106959).
- [22] P. Zhang, Z.Y. Yin, Y.F. Jin, Bayesian neural network-based uncertainty modelling: application to soil compressibility and undrained shear strength prediction, *Can. Geotech. J.* 59 (2022) 546–557, doi:[10.1139/cgj-2020-0751](https://doi.org/10.1139/cgj-2020-0751).
- [23] P. Zhang, Z.Y. Yin, Y.Y. Zheng, F.P. Gao, A LSTM surrogate modelling approach for caisson foundations, *Ocean Eng.* 204 (2020) 107263, doi:[10.1016/j.oceaneng.2020.107263](https://doi.org/10.1016/j.oceaneng.2020.107263).
- [24] S.L. Shen, N. Zhang, A. Zhou, Z.Y. Yin, Enhancement of neural networks with an alternative activation function tanhLU, *Expert Syst. Appl.* 199 (2022) 117181, doi:[10.1016/j.eswa.2022.117181](https://doi.org/10.1016/j.eswa.2022.117181).
- [25] K. Elbaz, S.L. Shen, H.M. Lyu, Prediction of disc cutter life during shield tunneling with AI via the incorporation of a genetic algorithm into a GMDH-type neural network, *Engineering* 7 (2) (2020) 238–251, doi:[10.1016/j.eng.2020.02.016](https://doi.org/10.1016/j.eng.2020.02.016).