Method Based on Support Vector Machine and Sequential Backward Selection for Seismic Liquefaction Potential Evaluation

Jianping LI[1,2], Runrun DONG[1], Jiansheng WANG[1] & Ling CHEN[3]

[1]College of Civil Engineering and Architecture, Henan University of Technology, Zhengzhou 450001, China
[2]School of Architectural Design and Engineering, Hebi Polytechnic ,Hebi 458030,China
[3]Audit Division , Hebi Polytechnic ,Hebi 458030,China

Abstract

[Object] In the paper, the support vector machine(SVM) is utilized to evaluate the earthquake-induced site liquefaction potential, and an optimization algorithm based on cross validation and sequential backward selection(SBS) is proposed to improve the generalization ability of the classifier for seismic liquefaction potential evaluation(SLPE).

[Methods] Usually, the accuracy of SLPE using the SVM varies greatly when the training dataset and test dataset change, so the classifier is not reliable enough in practice. Because cross validation is more convincing for evaluating the classifier performance in machine learning , the algorithm in the paper tries to reduce the maximum error of cross validation through adopting SBS to determine the input variables of the SVM. The performance of the classifier is assessed by the area under the curve (AUC) on the basis of confusion matrix.

[Results] As shown by data validation, the algorithm can reduce the maximum error of cross validation and the variation of accuracy in SLPE while maintaining good performance of the classifier.

[Conclusions] In conclusion, a method that can improve the reliability of SVMs for classification in SLPE is put forward in the paper.

Keywords: seismic liquefaction potential evaluation; support vector machine; sequential backward selection; cross validation

1. Introduction

The earthquake is a serious natural disaster, and it causes great losses to the society. The earthquake can lead to drastic changes of soil, which make building foundations lose strength and stiffness and then cause building damage. Seismic liquefaction is a main reason for instability and damage of buildings, so it is an important topic in seismic research and design of buildings [1]. In the earthquake，excess pore water pressure of soil is considered to be the main cause of foundation liquefaction[2]. Soil liquefaction has been found in many historical earthquakes[1, 3-5].

Macro phenomena of soil liquefaction such as sand boil, ground cracking, slope landslide, sinking and leaning of heavy buildings, floating of some light buildings are directly observable during earthquakes. But the liquefaction process of saturated soil is underground ,and its occurrence is related to certain geological conditions.

Researches in soil liquefaction mechanism are concerned by scholars both here and abroad, and the aim is to study the mechanical properties of saturated sandy or silty soil and to study the growth of pore water pressure under cyclic loading. Early studies showed that the liquefaction of

Corresponding author: Jianping LI(E-mail:EMAIL29@126.com)

saturated sand under earthquake is related to void ratio, confining pressure and cyclic stress[6]. In reference [2], experimental studies proved that the trigger condition for liquefaction is due to the large shear strain instead of the complete loss of effective stress resulting from pore pressure accumulation. On the basis of experiments , Adamidis and Madabhushi [7] studied the influence of soil drainage on pore pressures and shear stress-shear strain response in the earthquake. At the same time, the peak acceleration of ground motion plays a decisive role in the movement of soil particles, which affects the stress redistribution and liquefaction characteristics of soil [8]. In reference [9], the influence of seismic dynamic characteristics on foundation liquefaction was revealed to some extent.

The purpose of mechanism researches is to better understand the law of foundation liquefaction, so that the foundation liquefaction trend in earthquake can be accurately estimated. Liquefaction potential is the trend that excess pore water pressure produced in saturated sand or saturated silt during earthquake makes the effective shear strength of soil decrease or disappear, which results in sand boil or soil instability. The liquefaction potential evaluation methods based on experience of earthquake damage are in view of the investigations of liquefied soil layers in the past earthquakes, and the historical data are analogized to the new data for liquefaction discrimination, or the empirical relationship between the response of liquefied soil and various in-situ test indexes is utilized to evaluate the liquefaction potential. These methods mainly use in-situ test indexes such as standard penetration test(SPT), cone penetration test(CPT) and shear wave velocity test to assess the liquefaction potential[10-13]. On the basis of field test data, the liquefaction potential is evaluated according to some statistical criteria or empirical formulas.

In recent years ,owing to good ability in data processing and characteristics of nonparametric computation, machine learning has been widely used in classification[14]. In the field of seismic liquefaction potential evaluation (SLPE), several approaches of machine learning such as neural networks [15, 16], Bayesian networks [17], decision tree [18], random forest [19], support vector machines [20] have achieved good results and have showed their wide prospects of application.

However, the instability that the classification accuracy varies obviously on different training data sets exists in SLPE using machine learning. In most of existing SLPE studies, one-off validation is adopted. In one-off validation, the data set is split into training set and test set only one time, the training set is used to train the machine learning model, the test set is used to test the accuracy and generalization of the trained model. Perhaps, the model performs well on a certain training set and test set ,but it becomes bad when training set and test set change.

In machine learning, it is more convincing to adopt cross validation for testing the performance of a model[21-23]. In reference [24] ,5-fold cross validation was adopted ,and the means of cross validation were used to test the model. In fact, the calculation result of cross validation may vary obviously, so means are not appropriate for evaluating accuracy and generalization of the model. The support vector machine(SVM) is adopted using the data set including 226 cases [25] for SLPE, the accuracy of cross validation is listed in table 1 .

Table 1 Accuracy of cross validations with $k = 3,5,10$ for SLPE using SVM

k	Accuracy									
3	0.9600	0.9733	0.9200							
5	0.8889	1.0000	0.9778	0.9778	0.9333					
10	0.9130	0.8696	1.0000	1.0000	0.9565	0.9130	0.9565	1.0000	0.9565	0.8947

It is evident that accuracy varies greatly among runs in cross validation, and the variability is overlooked in one-off validation or testing model through means of cross validation. The model may achieve good performance on a certain training data set and test data set, but it may not achieve the same good performance on new data sets, so the model is unreliable in practice.

On the other hand, in previous SLPE studies, the features (i.e. input variables) are selected according to research conclusions on liquefaction mechanism or trial computation results [19, 26]. This kind of feature selection can be used to examine model accuracy on different feature sets, but it can not be used to reduce the variation of accuracy in cross validation .

In view of the above-mentioned problem, in this paper, a combination of the SVM and sequential backward selection(SBS) is employed. The SVM is used as binary classifier to assess seismic liquefaction potential，and SBS is adopted to select features on the basis of reducing the maximum error of cross validation. Calculation results show that the method proposed in this paper can reduce the fluctuation of classification accuracy, improve reliability and generalization performance of the model.

2. Theoretical basis

2.1 Support vector machine model

The SVM on the basis of statistical learning provides a new, efficient and novel model to improve generalization performance and can reach a global minimum. It can effectively solve problems with small samples[27, 28]. The SVM can fulfilled the tasks of classification, regression and distribution estimation [29].

Suppose the training vectors $x_i \in R^n$, i =1, $\cdots$, l , in two classes, and the label vector $y_i \in \{-1,1\}$, the ith case $d_i = (x_i, y_i)$, x_i is a n-dimensional vector of n features (i.e. model inputs), y_i is the class label(i.e. model outputs). The basic idea of SVMs is the concept of maximum margin hyper-planes. The classifier of SVM must choose the hyper-plane whose decision boundary has the largest margin, and it usually has smaller prediction errors than other hyper-planes with small margins. The basic formulations of the SVM for classification[29] are briefly described as follows:

As far as binary classification is concerned, the following primal optimization problem should be solved:

$$\min_{\omega,b,\xi} \frac{1}{2}\omega^T\omega + C\sum_{i=1}^{l}\xi_i \tag{1}$$

$$\text{subject to } y_i(\omega^T\phi(x_i)+b) \geq 1-\xi_i ,$$

$$\xi_i \geq 0 , i = 1 \ ,\cdots, l \ ,$$

where $\phi(x_i)$ maps x_i into a higher-dimensional space and $C>0$ is the regularization parameter. Due to the possible high dimensionality of the vector variable ω, usually the following dual problem is solved :

$$\min_{\alpha} \frac{1}{2}\alpha^T Q\alpha - e^T\alpha \tag{2}$$

$$\text{subject to } y^T\alpha = 0 ,$$

$$0 \le \alpha_i \le C, \quad i = 1, \cdots, l \quad ,$$

where $e = [1, \cdots, 1]^T$ is the vector of all ones , Q is an l by l positive semidefinite matrix, $Q_{ij} \equiv y_i y_j K(x_i, x_j)$,and $K(x_i, x_j) \equiv \phi(x_i)^T \phi(x_j)$ is the kernel function. In this paper, $K(x_i, x_j) = \exp(-\gamma \| x_i - x_j \|^2)$, $\gamma > 0$.

After problem (2) is solved, using the primal-dual relationship, the optimal ω satisfies

$$\omega = \sum_{i=1}^{l} y_i \alpha_i \phi(x_i) \tag{3}$$

and the decision function is

$$\operatorname{sgn}(\omega^T \phi(x) + b) = \operatorname{sgn}(\sum_{i=1}^{l} y_i \alpha_i K(x_i, x) + b) \tag{4}$$

2.2 Sequential backward selection

In addition to reducing the storage requirements and the computational cost, feature selection that excludes irrelevant variables can improve the generalization ability of the model[30]. A variety of search strategies, such as backward variable elimination, regularized elimination procedure, genetic algorithm (GA) [31]，wrapper approach [32]have been used for selecting input variables in machine learning. As a search algorithm ,the SBS tries to find the best feature subset by deleting features one by one in initial feature set ,and this feature selection method can improve the classification ability of the model [30]. The SBS is adopted in this paper, the feature selection process is described as follows:

With data of the form $D = \{(x_i, y_i), i = 1, \cdots, l\}$,where $x_i \in R^n$, $y_i \in \{-1, 1\}$. x_i is a n-dimensional vector of n features, y_i is the class label , s is the feature set of n features. The method consists of classifiers $f_0(x)$, $f_1(x)$ …, $f_m(x)$. $f_0(x)$ depends on the feature set s , $f_j(x)$ depends on only $n - j$ of the n features. There are C_{n-j}^{n} subsets with $n - j$ features, $f_j^k(x)$ is the classifier that depends on the feature set s_j^k , $s_j^k \in s$. If the error of the classifier $f_j^*(x)$ is smaller than all other $f_j(x)$, and its feature set is $s_j^* = \{x_{q_1}, \ldots, x_{q_{(n-j)}}\}$, then for each $x_{q_h} \in s_j^*$,let $s_{j+1}^h = s_j^* \setminus x_{q_h}$. If the error of $f_{j+1}^*(x)$ is smaller than all other $f_{j+1}(x)$ and $f_j^*(x)$, the classifier $f_{j+1}^*(x)$ based on s_{j+1}^* is taken as the optimal classifier. The above steps will repeat until the error of the classifier can't be further reduced.

2.3 Evaluation metrics

2.3.1 Cross validation

Cross validation is an important method to evaluate the generalization ability of models[23]. In k-fold cross validation, the data set is split randomly into k equal sized folds. In each run, k-1 folds are employed to constitute the training set for the classifier while the remaining fold is used as the test set for appraising the trained classifier through error or accuracy.

Usually, k is 10 or 30 in k-fold cross validation. As k increases, the number of cases in training set rises and a more robust classifier is acquired, but the test set gets smaller. At the same

time, if k is increased, the cost of training the classifier k times increases [33]. The 10-fold cross validation adopted in this paper is shown in Fig.1.

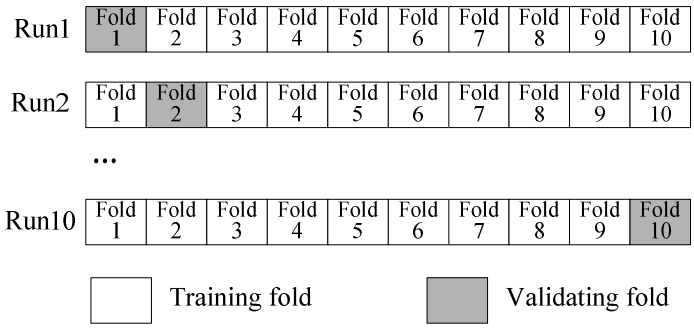

Fig.1 10-fold cross validation

2.3.2 Confusion matrix and ROC

The confusion matrix is a table describing the result of classification. SLPE is the problem of binary classification, and the evaluation result includes two categories: liquefied and non-liquefied. The confusion matrix can be shown as:

Table 2 Confusion matrix of SLPE

Actual Class	Predicted class	
	Liquefied (Positive)	Non-liquefied (Negative)
Liquefied (Positive)	TP	FN
Non-liquefied (Negative)	FP	TN

If a case is actually positive and the prediction is also positive, then it is a true positive case, the number of true positive cases is TP; if a positive case is predicted to be negative, it is a false negative case, the number of false negative cases is FN. With regard to a negative case, it is a true negative case if the prediction is also negative, the number of true negative cases is TN; the so-called false positive case is a negative case whose prediction is positive, the number of false positive cases is FP.

Several indexes ,such as error, accuracy, precision, true positive rate(TPR) and false positive rate(FPR) are defined to evaluate the classifier as follows[33]:

$$\text{error} = \frac{\text{FP+FN}}{\text{TP+TN+FP+FN}} \tag{5}$$

$$\text{accuracy} = \frac{\text{TP+TN}}{\text{TP+TN+FP+FN}} \tag{6}$$

$$\text{precision} = \frac{\text{TP}}{\text{TP+FP}} \tag{7}$$

$$\text{TPR} = \frac{\text{TP}}{\text{TP+FN}} \tag{8}$$

$$\text{FPR} = \frac{\text{FP}}{\text{FP+TN}} \tag{9}$$

In liquefaction problem, there are two types of errors, false positive cases and false negative cases. Obviously, the false negative cases are worse because they may lead to the neglect of potential danger. TPR measures what proportion of actual liquefaction cases to be discriminated, and FPR is the proportion of actual non-liquefaction cases being evaluated to be liquefiable.

Ideally, a classifier has a FPR of 0 and a TPR of 1. The ROC curve is a graphical plot developed by FPR on the x-axis and TPR on the y-axis, which presents a visualized view of a binary classifier's performance. The AUC is the area between the x-axis and the ROC curve, and the AUC can be a quantity value representing the classifier's expected performance comprehensively. Often, the AUC value is less than or equal to 1, and it is the bigger the better, so the AUC value of a classifier can be an index being compared with other classifiers [33].

3. The proposed method for SLPE

3.1 Data set

The data set [25] used in this paper includes 226 cases, there are 133 liquefied cases and 93 non-liquefied cases in the data set. These cases are obtained through CPT measurements at over 52 sites and field performance observations of 6 different earthquakes. The soils in these cases include clean sand, silty sand and silt mixtures (sandy and clayey silt), which are of the Holocene age.

As a machine learning approach, the SVM can find nonlinear quantitative relationships among data. Thus, the mapping from input to output is established. In this paper, the model is set up using the original data recorded in reference [25]. The feature set of the data set is $\{De, q_c, R_f, \sigma_v^{'}, \sigma_v, \alpha_{\max}, M_w\}$, where De is the depth of considered soil layer; q_c is the measured cone resistance; R_f is the friction ratio; $\sigma_v^{'}$ is the effective vertical stress; σ_v is the total vertical stress; $\alpha_{\max}$ is the peak horizontal ground surface acceleration; M_w is the moment magnitude.

3.2 Algorithm description

In the procedure of data preparation, scaling of input data is as follows:

$$x_{ij}^{'} = \frac{x_{ij} - \min\limits_{j}(x_{ij})}{\max\limits_{j}(x_{ij}) - \min\limits_{j}(x_{ij})} \tag{10}$$

Input data can be mapped to [0, 1] utilizing Equation 10, and the result variable $y_i \in \{-1, 1\}$, $y_i = -1$ denotes non-liquefied soil, $y_i = 1$ denotes liquefied soil.

The maximum error of cross validation is the optimization object of the proposed algorithm, and SBS is employed to select the optimal feature set. The proposed algorithm is as follows:

Algorithm

Input：Whole data set of n features, k (number of fold in cross validation)

Definition：Let F_n the whole set of the features，length-F_n the number of features in F_n, F_i a subset of F_n

Initialization: $F_i = F_n$

for i =0 to length- F_n

Let length- F_i the number of features in F_i

for j =1 to length- F_i

if $i \neq 0$

Let $f_j = j$ th feature of F_i

Let $F_j = F_i \setminus f_j$

else

$F_j = F_i$

end if

The data set is divided into k folds

for s =1 to k

sth fold is used for appraising the SVM classifier while the remaining folds are for training the classifier with the features set of F_j

end for

Then there are k errors while cross validation is completed, let err_{ij} the maximum of the k errors and record err_{ij}

if i =0

Break

end if

end for

Let $err_i = \min_j(err_{ij})$, F_i the features set corresponding to err_i

if $i \neq 0$ & $err_i > err_{i-1}$

Return F_{i-1}

Finish the computation

end if

end for

end Algorithm

4. Results and discussion

In this paper, the maximum error of cross validation is optimized through feature selection. So the classifier can achieve higher accuracy and better generalization, and the optimized model is more reliable in practice.

The optimal feature set is obtained through the proposed algorithm, it is $\{De, q_c, R_f, \sigma_v^{'}, \alpha_{\max}, M_w\}$,and σ_v is excluded from the optimal feature set.

There is the formula as

$$\sigma_v = \sigma_v^{'} + u \tag{11}$$

where u is the pore water pressure.

As far as liquefaction mechanism is concerned, σ_v and $\sigma_v^{'}$ are linearly dependent, and stress

is gradually transferred from soil skeleton to water as the pore water pressure in saturated sand or saturated silt increases during earthquake. Soil liquefaction is the process that $\sigma_v^{'}$ gradually decreases and u gradually increases [34]. σ_v can't reflect the process of stress redistribution in soil, so it is weakly related to soil liquefaction.

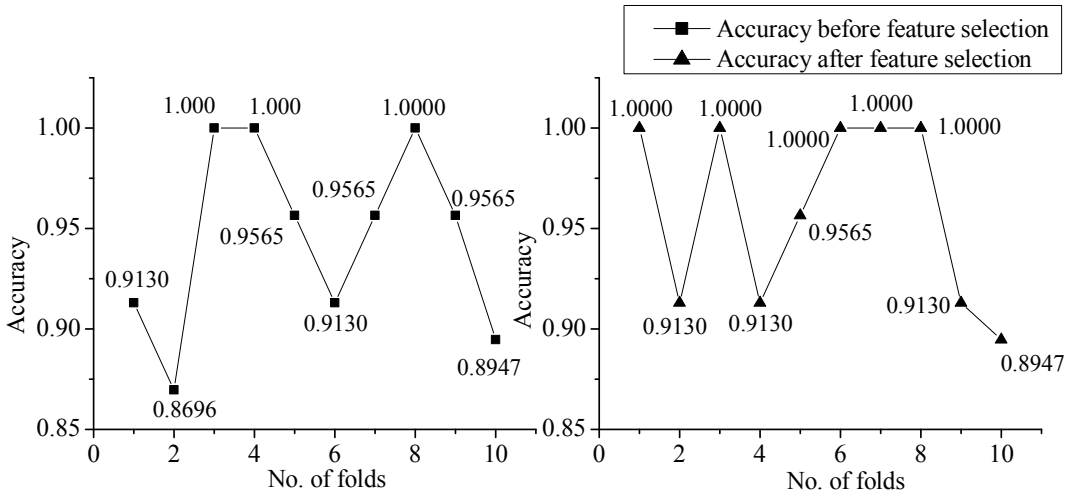

Fig.2 Accuracy of cross validation before and after feature selection

Fig.2 illustrate the accuracy of cross validation before and after feature selection. In 10-fold cross validation,3 folds are with the accuracy of 100% before feature selection, by contrast ,5 folds are with the accuracy of 100% after feature selection.The maximum error, minimum accuracy before and after feature selection are listed in table 3.

Table 3 Comparison before and after feature selection

	Maximum error error	Minimum accuracy
Before feature selection	0.1304	0.8696
After feature selection	0.1053	0.8947

It is clear that the accuracy and the generalization ability of the classifier are improved, the variation of calculation accuracy is reduced, and the calculation result is more reliable.

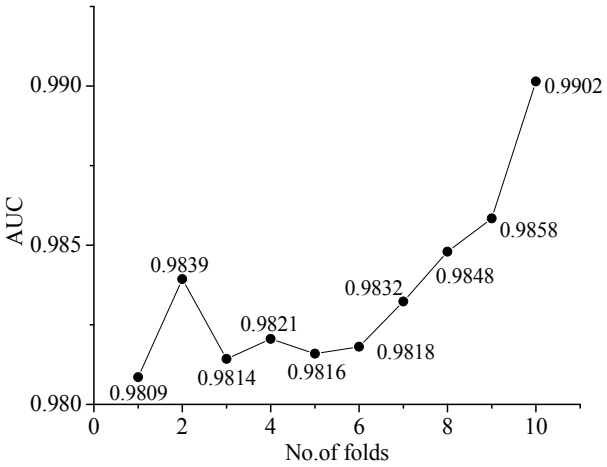

Fig.3 AUC of 10-fold cross validation

According to [35]，when AUC>0.9,the classifier is of outstanding discrimination. 10-fold cross validation is adopted in this paper, and the classifier after feature selection is trained by different training set. As shown in Fig.3, AUC of each fold in cross validation is more than 0.9, so the classifier after feature selection is of very good performance.

5. Conclusion

Machine learning technology has been applied to some extent to the SLPE, which is based on mass of seismic damage investigation data. How to find indicators closely related to soil liquefaction from mass of data is a subject needing further studies. On the other hand, machine learning is sensitive to data due to the prediction result variation caused by the change of training set. The sensibility limits the application of machine learning to SLPE. The object of the method mentioned in this paper is to reduce the sensibility of machine learning by means of SBS. Through the proposed algorithm, the maximum error of cross validation is reduced, the adaptability and generalization of the model are improved. Thus a new method that can improve reliability of the classifier in SLPE is put forward in this paper. However, the optimal feature set in this paper is only applicable to the data set adopted by this paper, so the cross validation and feature selection should be conducted again if there are new seismic liquefaction data.

References

[1]Y. Huang and M. Yu. Review of soil liquefaction characteristics during major earthquakes of the twenty-first century.Natural Hazards,2013, Vol. 65(3), p 2375-2384.
[2]D. Porcino and V. Diano. Laboratory study on pore pressure generation and liquefaction of low-plasticity silty sandy soils during the 2012 earthquake in Italian. Journal of Geotechnical and Geoenvironmental Engineering,2016, Vol. 142(10).
[3]M. Cubrinovski, J.D. Bray, C. de la Torre, M.J. Olsen, B.A. Bradley, G. Chiaro, E. Stocks, and L. Wotherspoon.Liquefaction effects and associated damages observed at the Wellington centreport from the 2016 Kaikoura earthquake.Bulletin of the New Zealand Society for Earthquake Engineering,2017, Vol. 50(2),p 152-173.
[4]L.W. Chen, X.M. Yuan, Z.Z. Cao, L.Q. Hou, R. Sun, L. Dong, W.M. Wang, F.C. Meng, and H.J. Chen.Liquefaction macrophenomena in the great Wenchuan earthquake. Earthquake Engineering and Engineering Vibration,2009, Vol. 8(2), p 219-229.
[5]R. Verdugo and J. Gonzalez.Liquefaction-induced ground damages during the 2010 Chile earthquake. Soil Dynamics and Earthquake Engineering,2015, Vol. 79, p 280-295.
[6]H.B. Seed. Liquefaction of saturated sands during cyclic loading. Journal of the Soil Mechanics and Foundation Engineering Division,1966, Vol. 92(SM6), p 105-134.
[7]O. Adamidis and S.P.G. Madabhushi.Experimental investigation of drainage during earthquake-induced liquefaction.Geotechnique,2018, Vol. 68(8), p 655-665.
[8]C. Clement, R. Toussaint, M. Stojanova and E. Aharonov.Sinking during earthquakes: Critical acceleration criteria control drained soil liquefaction.Physical Review E,2018, Vol. 97(2).
[9]S. Unjoh, M. Kaneko, S. Kataoka, K. Nagaya and K. Matsuoka. Effect of earthquake ground motions on soil liquefaction.Soils and Foundations,2012, Vol. 52(5), p 830-841.
[10]I. Akkaya, A. Ozvan, M. Akin, M.K. Akin and U. Ovun. Comparison of SPT and V-s-based liquefaction analyses: a case study in Ercis (Van, Turkey).Act a Geophysica,2018, Vol. 66(1), p 21-38.
[11]A. Isik, N. Unsal, A. Gurbuz and E. Sisman. Assessment of liquefaction potential of fethiye based on spt and shear wave velocity.Journal of the Faculty of Engineering and Architecture of Gazi University,2016, Vol. 31(4), p 1027-1037.
[12]X.L. Zhao and G.J. Cai.SPT-CPT correlation and its application for liquefaction evaluation in China. Marine Georesources & Geotechnology,2015, Vol. 33(3), p 272-281.
[13]M.W. Seo, S.M. Olson, C.G. Sun and M.H. Oh. Evaluation of liquefaction potential index along western coast of South Korea using SPT and CPT.Marine Georesources &

Geotechnology,2012, Vol. 30(3), p 234-260.
[14]A.E. Maxwell, T.A. Warner and F. Fang.Implementation of machine-learning classification in remote sensing: an applied review.International Journal of Remote Sensing,2018, Vol. 39(9), p 2784-2817.
[15]K. Venkatesh, V. Kumar and R.P. Tiwari.Appraisal of liquefaction potential using neural network and neuro fuzzy approach. Applied Artificial Intelligence,2013, Vol. 27(8), p 700-720.
[16]X.H. Xue and E.L. Liu.Seismic liquefaction potential assessed by neural networks. Environmental Earth Sciences,2017, Vol. 76(5), p 1-15.
[17]X.W. Tang, X. Bai, J.L. Hu and J.N. Qiu.Assessment of liquefaction-induced hazards using bayesian networks based on standard penetration test data.Natural Hazards and Earth System Sciences,2018, Vol. 18(5), p 1451-1468.
[18]A.H. Gandomi, M.M. Fridline and D.A. Roke.Decision tree approach for soil liquefaction assessment.Scientific World Journal,2013, Vol.2013,p1-8.
[19]V.R. Kohestani, M. Hassanlourad and A. Ardakani.Evaluation of liquefaction potential based on CPT data using random forest.Natural Hazards,2015, Vol. 79(2), p 1079-1089.
[20]X.H. Xue and X.G. Yang.Seismic liquefaction potential assessed by support vector machines approaches.Bulletin of Engineering Geology and the Environment,2016, Vol. 75(1),p 153-162.
[21]M. Fernandez-Delgado, E. Cernadas, S. Barro and D. Amorim.Do we need hundreds of classifiers to solve real world classification problems? Journal of Machine Learning Research,2014, Vol. 15, p 3133-3181.
[22]T.T. Wong and N.Y. Yang.Dependency analysis of accuracy estimates in k-fold cross validation.Ieee Transactions on Knowledge and Data Engineering,2017, Vol. 29(11),p 2417-2427.
[23]J.D. Rodriguez, A. Perez and J.A. Lozano.Sensitivity analysis of k-fold cross validation in prediction error estimation.IEEE Transactions on Pattern Analysis and Machine Intelligence,2010, Vol. 32(3), p 569-575.
[24]J.L. Hu, X.W. Tang and J.N. Qiu.Assessment of seismic liquefaction potential based on bayesian network constructed from domain knowledge and history data.Soil Dynamics and Earthquake Engineering,2016, Vol. 89, p 49-60.
[25]C.H. Juang, H. Yuan, D.H. Lee and P.S. Lin.Simplified cone penetration test-based method for evaluating liquefaction resistance of soils.Journal of Geotechnical & Geoenvironmental Engineering,2003, Vol. 129(1),p 66-80.
[26]P. Samui.Least square support vector machine and relevance vector machine for evaluating seismic liquefaction potential using SPT.Natural Hazards,2011, Vol. 59(2),p 811-822.
[27]P. Samui.Support vector machine applied to settlement of shallow foundations on cohesionless soils.Computers and Geotechnics,2008, Vol. 35(3), p 419-427.
[28]Y. Zhou, W. Su, L. Ding, H. Luo and P.E.D. Love.Predicting safety risks in deep foundation pits in subway infrastructure projects: Support vector machine approach. Journal of Computing in Civil Engineering,2017, Vol. 31(5).
[29]C.C. Chang and C.J. Lin.LIBSVM: A library for support vector machines.ACM Transactions on Intelligent Systems and Technology,2011, Vol. 2(3),p27:1-27:27.
[30]E. Romero and J.M. Sopena.Performing feature selection with multilayer perceptrons.IEEE Transactions on Neural Networks,2008, Vol. 19(3), p 431-441.
[31]M.M. Rahman, S.A. Imtiaz and K. Hawboldt.A hybrid input variable selection method for building soft sensor from correlated process variables.Chemometrics and Intelligent Laboratory Systems,2016, Vol. 157, p 67-77.
[32]L.H. Chen and C.J. Chang.Approaches to select suitable subset of explanatory variables for establishing fuzzy regression models.Journal of Intelligent & Fuzzy Systems,2018, Vol. 34(1), p 437-457.
[33]E. Alpaydin. Introduction to machine learning, 3rd ed.. Cambridge, Massachusetts: MIT Press, 2014.
[34]I.M. Idriss and R.W. Boulanger.Soil liquefaction during earthquakes.Oakland, California: Earthquake Engineering Research Institute, 2008.
[35]D.W. Hosmer, S. Lemeshow and R.X. Sturdivant, Applied logistic regression,3rd ed.. Hoboken, New Jersey: John Wiley & Sons, Inc, 2013.