

RESEARCHING THE APPLICABILITY OF VOTING-BASED ENSEMBLE MODELS FOR PREDICTING BRIDGE SCOUR DEPTH NEAR PIERS

NGHIÊN CỨU KHẢ NĂNG ỨNG DỤNG CỦA CÁC MÔ HÌNH VOTING TRONG VIỆC DỰ BÁO CHIỀU SÂU XÓI LỖ CỦA TRỤ CẦU

Vo Duyen Anh Huynh¹, Anh-Duc Pham², Thi Thu Ha Truong¹

¹Department of Civil Engineering; huynhvoduyenanh88@gmail.com, trttha@dct.udn.vn

²Faculty of Project Management; paduc@dut.udn.vn

Abstract - This paper proposes an ensemble model, namely voting to predict the bridge scour depth near piers. The voting model is constructed from four well-known individual learning classifiers, including artificial neural networks (ANN), support vector regression (SVR), classification and regression tree (CART) and linear regression (LR). The performance of the proposed model is evaluated by a reported dataset in the literature. The predictive accuracy of the voting model is compared with those obtained by four individual learning classifiers. Analytical results show that the voting model is superior to other comparative models in predicting the scour depth near piers. Particularly, the ANN+CART+LR is the best voting model that achieves the lowest synthesis index. The findings of this paper, therefore, provide a potential method for civil engineers in designing safe and cost-effective bridge structures.

Key words - Scour depth; bridge piers; individual learning classifiers; voting model; predictive accuracy.

1. Introduction

Bridges are essential structures that connect people to economic activities. A significant problem for the safety of bridges is bridge-pier scour [1, 2]. Local scour is defined as the removal of bed material from around piers, abutments, spurs, and embankments. This complex phenomenon is very harmful to the hydraulic structure itself and reduces service life of bridges. Therefore, an accurate and reliable estimate of local scour depths around bridge piers is vital in the hydraulic design of bridges [3].

The estimate of local scour depth at piers has attracted considerable research interest over years. Most investigations result in empirical equations that are based on regression models. In studies [4-6], the relationship between equilibrium scour depth near piers (d_s) and the decision variables is summarized as $d_s = f(\mu, V, V_C, h, g, \rho, D, d_{50})$; where μ is the dynamic viscosity of fluid; V is the average velocity of the approach flow, V_C is the critical velocity; h is the approach flow depth; g is the gravitational acceleration; ρ is the fluid density; D is the pier diameter; and d_{50} is the grain size.

The HEC-18 equation in [1] presents maximum scour depths around piers as $\frac{y_s}{y_1} = 2.0K_1K_2K_3\left(\frac{a}{y_1}\right)^{0.65}Fr_1^{0.43}$

where y_s is the scour depth; y_1 denotes the flow depth directly upstream of the pier; K_1 denotes correction factor for pier nose shape; K_2 denotes correction factor for angle of attack of flow; K_3 denotes correction factor for bed condition; a denotes the pier width; and Fr_1 denotes Froude Number directly upstream of the pier.

Tóm tắt - Bài báo này đề xuất một mô hình kết hợp tên là voting để dự báo chiều sâu xói lở của trụ cầu. Mô hình voting được xây dựng từ bốn mô hình học máy đơn lẻ nổi tiếng, bao gồm mạng nơ ron nhân tạo (ANN), máy học vector hỗ trợ (SVR), cây phân loại và hồi quy (CART) và hồi quy tuyến tính (LR). Sự thể hiện của mô hình voting được đánh giá bởi số liệu đã được công bố. Độ chính xác dự báo của mô hình voting được so sánh với bốn mô hình học máy đơn lẻ. Kết quả phân tích cho thấy, mô hình voting tốt hơn hẳn so với các mô hình so sánh khác trong dự báo chiều sâu xói lở của trụ cầu. Đặc biệt, ANN+CART+LR là mô hình voting tốt nhất, mô hình này đạt sai số dự báo nhỏ nhất. Kết quả của bài báo này, vì vậy, cung cấp một phương pháp tiềm năng cho các kỹ sư dân dụng trong việc thiết kế cấu trúc công trình cầu an toàn và hiệu quả về chi phí.

Từ khóa - Chiều sâu xói lở; trụ cầu; mô hình học máy đơn lẻ; mô hình voting; độ chính xác dự báo.

However, these empirical methods that focused on special parameters could lead to the prediction of scour depth with low accuracy [7]. Recently, artificial intelligence (AI)-based approaches have been successfully applied in solving various problems in engineering. For instances, Lee et al. (2007) [8] applied back-propagation neural network (BPN) - a common type of neural networks - to estimate the scour depth around bridge piers. Experimental results indicated that the BPN model outperformed empirical methods in predicting the scour depth. Particularly, support vector regression (SVR) – a version regression of support vector machines [9] – has been adopted to predict the scour depth. Pal et al. [3] investigated the ability of the SVR in modeling the local scour around bridge piers using field data. Analytical results showed a better forecasting performance of the SVR compared to four empirical relations, BPN and generalized regression neural network.

Single AI-based models (e.g, ANN, SVR) have their own drawbacks and such models could not effectively solve all complex problems in real worlds. A combination of single AI-based models to take advantage of each model is essential. Voting is an ensemble model that combines multiple prediction models (or learning classifiers) to improve the performance of single (or baseline) models. In AI-based applications, artificial neural networks (ANN), SVR, classification and regression tree (CART), and linear regression (LR) are well-known individual learning classifiers. Moreover, no study in literature has compared such individual and ensemble learning techniques for predicting scour depth near piers.

This study, therefore, investigates the predictive performance of four individual models and voting-based

ensemble models in predicting scour depths. Experimental data collected from published studies is used to evaluate the performance of the considered models. The predictive performances of each model are compared in terms of root mean squared error (*RMSE*), mean absolute error (*MAE*), and mean absolute percentage error (*MAPE*) via a synthesis index (*SI*).

The rest of this paper is organized as follows. Section 2 elucidates individual AI-based models, voting model and performance evaluation methods. Experimental settings are presented in Section 3. Section 4 shows analytical results and discussion. Finally, the conclusions and future study directions are shown in Section 5.

2. Methodologies

2.1. Individual artificial intelligence-based models

2.1.1. Artificial neural networks

Artificial neural networks (ANN) have proven effective for solving nonlinear models because of their universal approximation property. Multilayer Perceptron is a feedforward neural network that maps sets of input data onto a set of appropriate outputs. This model consists of an input layer containing a set of sensory input nodes, one or more hidden layers containing computation nodes, and an output layer containing one computation node.

One of the most widely used training algorithms is the backpropagation algorithm (BP), which adjusts connection weights and bias values during training. The Eq. (1) expresses an activated neuron in a hidden output layer.

$$net_j = \sum w_{ji}x_i \text{ and } y_j = f(net_j) \quad (1)$$

where net_j denotes the activation of j th neuron; i denotes the set of neurons in the preceding layer; w_{ji} denotes the weight of the connection between neuron j and neuron i ; x_i denotes the output of neuron i ; and y_j denotes the sigmoid or logistic transfer function.

$$f(net_j) = \frac{1}{1 + e^{-\lambda net_j}} \quad (2)$$

where λ controls the function gradient.

The formula for training and updating weights w_{ji} in each cycle h is determined as in Eq. (3).

$$w_{ji}(h) = w_{ji}(h-1) + \Delta_{ji}(h) \quad (3)$$

Here, $\Delta_{ji}(h)$ is the change.

$$\Delta_{ji}(h) = \eta \delta_{pi} \chi_{pi} + \alpha \Delta w_{ji}(h-1) \quad (4)$$

where η is the learning rate parameter; δ_{pi} is the propagated error; χ_{pi} is the output of neuron i for record p ; α is the momentum parameter; and $\Delta w_{ji}(h-1)$ is the change in w_{ji} in the previous cycle.

2.1.2. Support vector machines for regression

Support vector regression (SVR) is a regression version of support vector machines. Developed by Vapnik (1995) [9], SVR seeks to minimize an upper bound of the generalization error than minimize the empirical error in neural networks. In this method, an ε -insensitive loss function is used to nonlinearly map the input space into a

high-dimensional feature space, and then run linear regression in the output space. The SVR can thus be formulated by simplifying the following function.

$$\min \frac{1}{2} \|\omega\|^2 + C \sum_{i=1}^n (\xi_i + \xi_i^*) \quad (5)$$

subject to $y_i - f(x_i, \omega) \leq \varepsilon + \xi_i^*$; $f(x_i, \omega) - y_i \leq \varepsilon + \xi_i$;
 $\xi_i, \xi_i^* \geq 0, i = 1, \dots, n$

where ω is the parameter of the linear approximator; $C \geq 0$ is a regularization constant that represents the trade-off between the empirical error and the flatness of the function; ξ and ξ^* are non-negative slack variables; x_i is input patterns; y_i is prediction labels; and n is the sample size.

This optimization problem can be transformed into a dual problem, which is solved by

$$f(x) = \sum_{i=1}^{n_{SV}} (\alpha_i - \alpha_i^*) K(x, x_i) \quad (6)$$

subject to $0 \leq \alpha_i^* \leq C$; $0 \leq \alpha_i \leq C$

where n_{SV} is the number of support vectors and $K(x, x_i)$ is the kernel function. During training, kernel functions are used to identify support vectors along the function surface. In highly nonlinear spaces, a radial basis function (RBF) kernel usually yields better results than other kernel [10].

2.1.3. Classification and regression tree

Introduced by Breiman et al. [11], the CART refers to decision tree algorithms that can be used for classification or regression predictive modeling problems. Classification trees are designed for dependent variables that take a finite number of unordered values, with prediction error measured in terms of misclassification cost. Regression trees are for dependent variables that take continuous or ordered discrete values, with prediction error typically measured by the squared difference between the observed and predicted values [12]. A learning tree can be optimized by using a learning data set to prune the saturated tree and select among the obtained sequence of nested trees [11].

2.1.4. Linear regression

Linear regression is an advanced version of the simple regression model, which determines the relationship between a scalar dependent variable and two or more explanatory variables. This model specifies that an appropriate function of the fitted probability of the event is a linear function of the observed values of the available explanatory variables. Eq. (7) denotes a general formula of linear regression.

$$Y = \beta_o + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_n x_n + \varepsilon \quad (7)$$

where Y is the dependent variable; β_o is the constant term in the model; β_i is a regression coefficient ($i=1,2,\dots,n$); ε is an error term; and x_i represents explanatory variables ($i=1,2,\dots,n$). In this study, Y denotes the resilient modulus of subgrade soils, and x_i denotes i^{th} attribute of resilient modulus.

2.2. Voting model

Voting is a method of combining multiple classifiers [13, 14]. In prediction problems, the outputs of the individual

classifiers are pooled. The class with the largest number of votes is then selected as the final classification decision. Generally, the numerical output can be determined by different combinations of probability estimates. Figure 1 presents the architecture of a voting ensemble model.

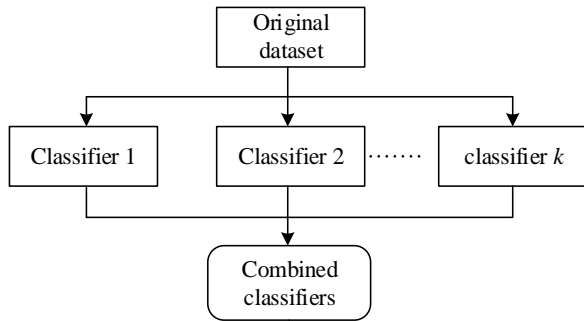


Figure 1. Voting scheme

This study obtains eleven ensemble classifiers consisting of two to four different individual classifiers. The two-classifier ensembles are ANNs+CART, ANNs+SVMs, ANNs+LR, CART+SVR, CART+LR, and SVR+LR. The three-classifier ensembles include ANNs+CART+SVR, ANNs+CART+LR, CART+SVR+LR, and ANNs+SVR+LR. One four-classifier ensemble is ANNs+CART+SVR+LR.

2.3. Performance evaluation methods

To assess the prediction accuracy of proposed models, some performance measures are used. They include *RMSE*, *MAE*, *MAPE*, and *SI*. The corresponding equations are as follows.

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y' - y)^2} \quad (8)$$

where y is the actual value; y' is the predicted value; and n is the number of sample data.

$$MAPE = \frac{1}{n} \sum_{i=1}^n \left| \frac{y - y'}{y} \right| \quad (9)$$

$$MAE = \frac{1}{n} \sum_{i=1}^n |y - y'| \quad (10)$$

$$SI = \frac{1}{m} \sum_{i=1}^m \frac{P_i - P_{\min,i}}{P_{\max,i} - P_{\min,i}} \quad (11)$$

where m is the number of performance measures (in this study, $m = 3$) and P_i is the performance measure i^{th} . The *SI* ranges from 0 to 1 and *SI* value close to 0 indicates a highly accurate predictive model.

3. Experimental setting

3.1. Data description

This study employs the voting-based ensemble model to increase predictive accuracy when predicting bridge scour depth near pier. The total dataset including 151 samples is collected from published studies [4-6]. The predictor variables include pier diameter, grain diameter mean, approach flow depth, flow velocity, and critical velocity. The response is the pier scour near piers. Statistical parameters are summarized in Table 1.

Table 1. Statistical parameters for the dataset

Parameter	Unit	Min	Average	Max
Input				
Pier diameter	mm	16.00	84.20	200.00
Grain diameter mean	mm	0.30	1.50	7.80
Approach flow depth	mm	20.00	256.30	600.00
Flow velocity	m/s	0.20	0.60	1.30
Critical velocity	m/s	0.30	0.50	1.30
Output				
Pier scour	mm	4.00	132.50	615.40

3.2. Model construction

The total dataset is separated into a training data and a test data. The training data of 136 samples (90 percent of the total dataset) is used to the model. The test data of 15 samples (10 percent of the total dataset) is then used to evaluate the trained model.

Such selection of training and test data may lead to the bias. In this paper, a k -fold cross validation algorithm is employed to minimize bias associated with randomly sampled training and test data. Kohavi (1995) showed that 10 folds are optimal [15]. Thus, 10-fold cross validation is adopted to validate the model performance. In each round, nine folds are used to train the model and the remaining fold is used for model training. This procedure is repeatedly performed in 10 times with a distinct fold as test data. Figure 2 describes a flowchart for training and testing individual AI and voting models.

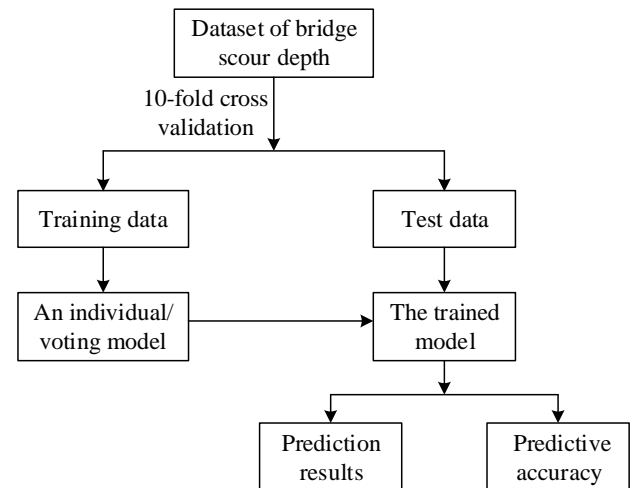


Figure 2. Flowchart of the model training and testing

The four individual AI-based models used to construct the voting model are ANN, SVR, CART, and LR. The efficacy of these models are investigated using the open-source software WEKA [16]. Table 2 shows parameter set as default to develop these individual and voting models.

Table 2. Parameter setting of models in WEKA

Model	Parameter	Setting
ANN	Hidden layer	3
	Leaning rate	0.3
	Momentum	0.2

	Training/time	500
	Seed	0
SVR	C	1.0
	Kernel	RBF
CART	Initial count	0.0
	Max depth	-1
	MinNum	2.0
	MinVarianceProp	0.001
	NoPruning	False
	NumFolds	3
	Seed	1
LR	Attribute Selection Method	M5 method
	Eliminate Colinear Attributes	True
	Debug	False
	Ridge	1.0E-8
Voting	Classifiers	2-4 weka.classifiers. Classifier
	Combination Rule	Average of probabilities
	Seed	1

4. Analytical results and discussion

In this paper, the performance measures based on test data are analyzed since the predictive ability of a model is confirmed in respecting with the testing phase. Table 3 compares the predictive performance of four baseline models and eleven voting-based ensemble models. The SI ranks the performance of these models. Of four individual models, the ANN shows the best performance with the smallest values of RMSE and MAE (37.26 mm and 25.95 mm, respectively). Following the ANN are the LR, CART, and the SVR, respectively. Particularly, the CART yields the smallest value of MAPE of 25.20%

The SI in Table 3 indicates that all ensemble learning classifiers are superior to individual ones in predicting the scour depth. In spite of the outperformance of the ANN among baseline models, its predictive ability is worse than all eleven voting models except for the SVR+LR. Among fifteen learning classifiers, the ANN+CART+LR voting model is the best for predicting the scour depth. It yields smallest values of RMSE and MAE with 32.38mm and 22.02mm, respectively. The analytical in Table 3 also shows that the ANN+CART is the second best model with the lowest MAPE of 21.61%. Especially, the voting model formed by four individual classifiers - ANN+CART +SVR+LR - ranks third among all comparative models in predictive ability.

Table 3. Predictive performance comparison of individual models and voting model

Model	RMSE (mm)	MAE (mm)	MAPE (%)	SI (Ranking)
Baseline model				
ANN	37.26	25.95	27.30	(11)
SVR	41.67	29.19	39.50	(15)
CART	39.89	27.70	25.20	(14)
LR	35.46	26.10	30.41	(12)

Voting-based ensemble model				
ANN+SVR	34.41	24.02	28.35	0.291 (8)
ANN+CART	33.47	23.04	21.61	0.087 (2)
ANN+LR	33.55	23.18	25.45	0.167 (5)
CART+SVR	36.20	25.24	29.16	0.427 (10)
CART+LR	34.24	23.73	24.32	0.197 (6)
SVR+LR	36.58	26.05	33.71	0.563 (13)
ANN+CART+SVR	32.85	23.04	25.30	0.133 (4)
ANN+CART+LR	32.38	22.02	22.34	0.014 (1)
CART+SVR+LR	34.44	24.20	28.27	0.299 (9)
ANN+SVR+LR	33.67	23.52	28.09	0.237 (7)
ANN+CART+SVR+LR	32.61	22.79	25.59	0.118 (3)

Note: Bold values denote the best performance measures in each type of AI model; (*) stands for model raking.

Figure 3 further compares the predictive performance of all learning classifiers in terms of MAPE. Obviously, the combination of three or more individual models achieves better performance than those obtained by one or two individual models.

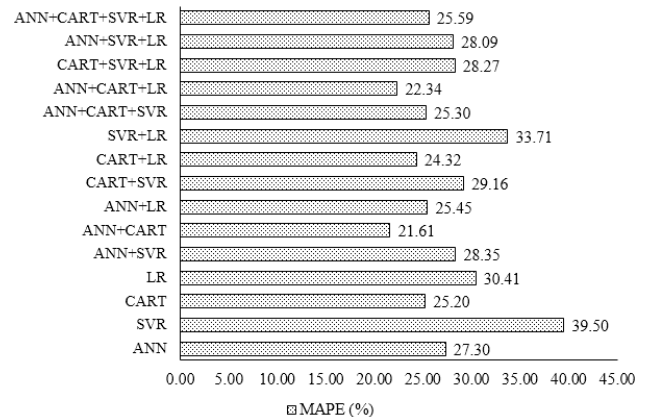


Figure 3. Performance comparison of all models in term of the MAPE

5. Conclusions

This paper investigates the predictive ability of various machine learning techniques in predicting bridge scour depth near piers. The ensemble learning classifier – voting – is constructed from four well-known machine learners, including ANN, SVR, CART, and LR. The efficacy of individual and voting models is validated by a published laboratory dataset. The whole dataset is divided into training data and test data using 10-fold cross validation algorithm. Four performance measures, including RMSE, MAE, MAPE, and SI are utilized to compare the predictive performance of examined models.

The analytical results show the outperformance of voting-based ensemble models to individual learning classifiers. Particularly, the voting-based ensemble model significantly improves predictive accuracy of individual learning classifiers in predicting bridge scour depth except for the SVR+LR. Overall, the voting-based ANN+CART+LR is the best learning classifier that achieves lowest error rates in terms of RMSE (32.38mm) and MAE (22.02mm).

The contribution of this paper is to investigate the voting-based ensemble models for predicting bridge scour depth near piers. The findings of this paper indicate a potential applicability of the voting-based ensemble model in solving complex nonlinear problems in civil engineering. Future studies, therefore, should comprehensively investigate the performance of voting model via various datasets.

REFERENCES

- [1] L.A. Arneson, L.W. Zevenbergen, P.F. Lagasse, P.E. Clopper, Evaluating scour at bridges, Federal Highway Administration, Washington, DC, 2012.
- [2] H.M. Azamathulla, A.A. Ghani, N.A. Zakaria, A. Guven, Genetic Programming to Predict Bridge Pier Scour, *Journal of Hydraulic Engineering* 136(3) (2010) 165-169.
- [3] M. Pal, N.K. Singh, N.K. Tiwari, Support vector regression based modeling of pier scour using field data, *Engineering Applications of Artificial Intelligence* 24(5) (2011) 911-916.
- [4] B.W. Melville, Y.-M. Chiew, Time scale for local scour at bridge piers, *Journal of Hydraulic Engineering* 125(1) (1999) 59-65.
- [5] C.S. Lauchlan, B.W. Melville, Riprap Protection at Bridge Piers, *Journal of Hydraulic Engineering* 127(5) (2001) 412-418.
- [6] M.F. Mia, H. Nago, Design Method of Time-Dependent Local Scour at Circular Bridge Pier, *Journal of Hydraulic Engineering* 129(6) (2003) 420-427.
- [7] M. Najafzadeh, G.-A. Barani, H.M. Azamathulla, GMDH to predict scour depth around a pier in cohesive soils, *Applied Ocean Research* 40 (2013) 35-41.
- [8] T.L. Lee, D.S. Jeng, G.H. Zhang, J.H. Hong, Neural network modeling for estimation of scour depth around bridge piers, *Journal of Hydrodynamics*, Ser. B 19(3) (2007) 378-386.
- [9] V.N. Vapnik, *The nature of statistical learning theory*, Springer-Verlag, New York, 1995.
- [10] H. Wang, D. Hu, Comparison of SVM and LS-SVM for regression, International Conference on Neural Networks and Brain, *IEEE*, 2005, pp. 279-283.
- [11] L. Breiman, J.H. Friedman, R.A. Olshen, C.J. Stone, Classification and Regression Trees, Chapman and Hall/CRC, New York, 1984.
- [12] W.-Y. Loh, Classification and regression trees, *Wiley Interdisciplinary Reviews: Data Mining Knowledge Discovery*, 2011, pp. 14-23.
- [13] J. Kittler, M. Hatef, R.P.W. Duin, a.J. Matas, On Combining Classifiers, *IEEE transactions on pattern analysis and machine intelligence* 20(3) (1998) 226-239.
- [14] L.I. Kuncheva, *Combining Pattern Classifiers: Methods and Algorithms*, Wiley, New York City, United States, 2004.
- [15] K. R., A study of cross-validation and bootstrap for accuracy estimation and model selection, *International joint conference on artificial intelligence*, Morgan Kaufmann, 1995, pp. 1137-43.
- [16] Machine Learning Group at the University of Waikato. <http://www.cs.waikato.ac.nz/ml/weka/>.

(The Board of Editors received the paper on 02/10/2017, its review was completed on 15/10/2017)